ESCAP MEETING NO. 7 - 04/12/00

AGENDA

Kathleen P Zveare
04/10/2000 03:29 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC
cc:
Subject: 4/12 ESCAP Meeting

The April 12 ESCAP Meeting will be held in Rm. 2412/3 from 10:30-12. The agenda is as follows:

Post-Stratification - Donna Kostanich/Howard Hogan

ESCAP MEETING NO. 7 - 04/12/00

HANDOUTS

UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

April 19, 2000

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES #Q-24

MEMORANDUM FOR Howard Hogan
Chief, Decennial Statistical Studies Division

From: Donna Kostanich
Assistant Division Chief, Sampling and Estimation
Decennial Statistical Studies Division

Prepared by: Richard Griffin
Dawn Haines
Estimation Staff

Subject: Accuracy and Coverage Evaluation Survey: Final Post-stratification Plan for Dual System Estimation

I. INTRODUCTION

The goal of post-stratification is to group together people who have similar coverage by the census. A common assumption is that people who share similar housing, similar language, similar cultural attitudes, and similar education would also share similar census coverage. Tenure, race and ethnic origin often serve as a marker for these similarities.

This memorandum presents the final post-stratification plan for the Accuracy and Coverage Evaluation (A.C.E.) Survey including Puerto Rico. The plan for Census 2000 A.C.E. is summarized in Section III. The detailed definitions of the poststratification variables and the race and Hispanic origin groups are given in Sections IV. and V., respectively.

II. BACKGROUND

The 2000 A.C.E. is different from the 1990 Post Enumeration Survey (PES). The A.C.E. will have approximately twice the sample size of the PES. This larger sample size allows for the formation of more post-strata and more post-strata have the advantage of reducing correlation bias. Additionally in 2000 multiple responses to the race question will be permitted; whereas in 1990 only one race could be selected.

The 1990 PES had 357 post-strata defined by a cross-classification of 51 post-stratum groups by seven age/sex groups. The 357 design started with a cross-classification of seven variables: age, sex, race, Hispanic origin, tenure, urbanicity, and region. There were 840 cells in the cross-classification. Collapsing was necessary in order to produce post-strata with sufficient sample for reliable Dual System Estimation (DSE). The attachment shows the 51 post-stratum groups for the 1990 PES after collapsing and the seven age/sex groups.

Race and Hispanic origin were the most important variables. After collapsing, five race/Hispanic origin post-strata were maintained: Non-Hispanic White or Other, Black, Hispanic White or Other, Asian and Pacific Islander, and Reservation Indians. Off-reservation American Indians were placed in either the Non-Hispanic White or Other group or the Hispanic White or Other group depending on whether they were of Hispanic origin. Within each of these race/Hispanic origin post-strata, seven age/sex categories were maintained.

The other variables were collapsed in the following order: region, urbanicity, then tenure, if necessary. For American Indians residing on reservations, all these variables were collapsed. For Asian and Pacific Islanders, region and urbanicity were collapsed and tenure maintained. For the Black and Hispanic White or Other groups, region was collapsed for two levels of urbanicity. For Non-Hispanic White or Other, the full cross-classification of region, urbanicity and tenure were maintained.

The 1990 PES for Puerto Rico had 21 post-strata defined by a cross-classification of 3 Place Type categories and seven age/sex categories. The place types were central city areas in Metropolitan Statistical Areas, non-central cities in Metropolitan Statistical Areas, and areas outside of Metropolitan Statistical Areas. The seven age/sex categories were the same as those used for the U.S. These 1990 post-stratification groups for Puerto Rico are also given in the attachment.

III. CENSUS 2000 A.C.E. POST-STRATIFICATION PLAN

For the Census 2000 A.C.E. we will retain most of the 1990 PES post-stratification variables and we will include several additional variables. The 2000 A.C.E. post-strata will be defined by nine variables: age, sex, race, Hispanic origin, tenure, region, Metropolitan Statistical Area size, Type of Enumeration Area, and tract level return rate. The Metropolitan Statistical Area size variable is replacing the urbanicity variable which will not be available until the summer of 2001. Type of Enumeration Area and the tract return rate are two new features of the 2000 A.C.E. post-stratification. The mailout/mailback areas will be differentiated from other types of enumeration areas. Tracts will be classified by high or low return rate. Additionally, multiple responses to the race question will be reflected in the race and Hispanic origin groupings.

Table 1a shows the 64 post-stratum groups for the Census 2000 A.C.E.. Within each post-stratum group there will be seven age/sex groups (Table 1c). Thus, there is a maximum of $64 \times 7 = 448$ post-strata, and of course there will be fewer if further collapsing is necessary. The post-stratification plan was chosen to reduce correlation bias without having an adverse effect on the variance of the Dual System Estimator.

For the Census 2000 A.C.E. in Puerto Rico, post-strata will be defined by five variables: age, sex, tenure, Metropolitan Statistical Area, and tract-level return rate. The variable region is not applicable for Puerto Rico. Further, there is only one Type of Enumeration Area (Update/Leave) in Puerto Rico, so this variable is not utilized. Table 1b shows the 12 post-stratum groups used in the Puerto Rico Census 2000 A.C.E. Survey. Within each post-stratum group, the seven age/sex groups in Table 1c are utilized. Thus, there is a maximum of $12 \times 7 = 84$ post-strata, and of course there will be fewer if further collapsing is necessary.

Tables 1a and 1b show the 64 and 12 post-stratum groups for the U.S. and Puerto Rico, respectively. Table 1c presents the seven age/sex groups which are used for both the U.S. and Puerto Rico. Subsequent sections of this memorandum provide a detailed description of the post-stratification domains and variables, including any alternative definitions for Puerto Rico. An extensive explanation of the domains is presented in Section V.

Table 1a: Census 2000 A.C.E. - 64 Post-Stratum Groups (U.S.)

Race/Hispanic Origin Domain Number*		Tenure	MSA/TEA	High Return Rate				Low Return Rate			
				N	M	S	W	N	M	S	W
Domain 7 (Non-Hispanic White or "Some other race")		Owner	Large MSA MO/MB	1	2	3	4	5	6	7	8
			Medium MSA MO/MB	9	10	11	12	13	14	15	16
			Small MSA & Non-MSA MO/MB	17	18	19	20	21	22	23	24
			All Other TEAs	25	26	27	28	29	30	31	32
		Non-owner	Large MSA MO/MB	33				34			
			Medium MSA MO/MB	35				36			
			Small MSA & Non-MSA MO/MB	37				38			
			All Other TEAs	39				40			
Domain 4 (Non-Hispanic Black)		Owner	Large MSA MO/MB	41				42			
			Medium MSA MO/MB								
			Small MSA & Non-MSA MO/MB	43				44			
			All Other TEAs								
		Non-owner	Large MSA MO/MB	45				46			
			Medium MSA MO/MB								
			Small MSA & Non-MSA MO/MB	47				48			
			All Other TEAs								
Domain 3 (Hispanic)		Owner	Large MSA MO/MB	49				50			
			Medium MSA MO/MB								
			Small MSA & Non-MSA MO/MB	51				52			
			All Other TEAs								
		Non-owner	Large MSA MO/MB	53				54			
			Medium MSA MO/MB								
			Small MSA & Non-MSA MO/MB	55				56			
			All Other TEAs								
Domain 5 (Native Hawaiian or Pacific Islander)		Owner		57							
		Non-owner		58							
Domain 6 (Non-Hispanic Asian)		Owner		59							
		Non-owner		60							
American Indian or Alaska Native	**Domain 1** (On Reservation)	Owner		61							
		Non-owner		62							
	Domain 2 (Off Reservation)	Owner		63							
		Non-owner		64							

- For Census 2000 persons can self identify with more than one race group. For post-stratification, persons are included in a single Race/Hispanic Origin domain. This does not change a person's actual response and all persons will be tabulated based on their actual response in the census. An extensive explanation of the domains is presented in Section V.

Table 1b: Census 2000 A.C.E. - 12 Post-Stratum Groups (Puerto Rico)

Tenure	MSA	High Return Rate	Low Return Rate
Owner	San Juan CMSA	1	2
	Other MSA	3	4
	Non-MSA	5	6
Non-owner	San Juan CMSA	7	8
	Other MSA	9	10
	Non-MSA	11	12

Table 1c: Census 2000 A.C.E. - 7 Age/Sex Groups (U.S. & Puerto Rico)

	Male	Female
Under 18	A	
18 to 29	B	C
30 to 49	D	E
50+	F	G

Key:

Return Rate: Tract-level variable measuring the proportion of occupied housing units in the mailback universe which returned a census questionnaire. Low return rate tracts are those tracts whose return rate is less than or equal to the 25th percentile return rate.

MSA: Metropolitan Statistical Area or Consolidated Metropolitan Statistical Area, as defined by the Office of Management and Budget (OMB), will be referred to as MSA.

TEA: Type of Enumeration Area.

MO/MB: Mailout/Mailback Type of Enumeration Area.

N, M, S, W: Refers to region - Northeast, Midwest, South, West.

"Some other race": One of six possible major race categories obtained from the census questionnaire.

Further details on the variable definitions are included in the following sections.

IV. CENSUS 2000 POST-STRATIFICATION VARIABLES

A. Post-stratification Variables

A.C.E. post-stratification will use the following variables:

- Race/Hispanic Origin - seven categories (omitted for Puerto Rico)
- Age/Sex - seven categories
- Tenure - two categories
- Metropolitan Statistical Area (MSA) by Type of Enumeration Area (TEA) - four categories (three categories for Puerto Rico)
- Return Rate - two categories
- Region - four categories (omitted for Puerto Rico)

The seven Race/Hispanic Origin domains are:

- American Indian or Alaska Native on Reservations
- Off-Reservation American Indian or Alaska Native
- Hispanic
- Non-Hispanic Black
- Native Hawaiian or Pacific Islander
- Non-Hispanic Asian
- Non-Hispanic White or "Some other race"

See Section V. for further details on the Race/Hispanic Origin domains. Inclusion in a Race/Hispanic Origin domain is complicated as it depends on several variables and whether there are multiple race responses. In addition, inclusion in a Race/Hispanic Origin domain **does not** change a persons Race/Hispanic Origin response. All Census 2000 tabulations will be based on the actual responses. For example, a person who responds as American Indian on a reservation and Black will be placed in the first Race/Hispanic Origin domain (Group 1) for post-stratification purposes but will be tabulated in the census as American Indian/Black.

The seven Age/Sex categories are:

- Under 18
- 18 - 29 Male
- 18 - 29 Female
- 30 - 49 Male
- 30 - 49 Female
- 50+ Male
- 50+ Female

The two Tenure categories are:

- Owner

- Non-owner

The four MSA/TEA categories are:

- Large MSA Mailout/ Mailback (MO/MB)
- Medium MSA MO/MB
- Small MSA or Non-MSA MO/MB
- All other TEAs

MSA/CMSA FIPS codes, as defined by the Office of Management and Budget (OMB), will be used for post-stratification. For simplification, MSA/CMSA will herein be referred to as MSA. Large MSA consists of the ten largest MSAs based on unadjusted, Census 2000 total population counts including the population in Group Quarters. Medium MSAs are those (besides the largest 10) which have at least 500,000 total population. Small MSAs are those with a total population size strictly less than 500,000. For post-stratification purposes, MO/MB areas are contrasted with the non-MO/MB areas.

For Puerto Rico there are three MSA categories. The TEA portion of this variable is nonexistent since all of Puerto Rico is Update/Leave. The three MSA categories are:

- San Juan CMSA (San Juan-Caguas-Arecibo CMSA)
- Other MSA (Aguadilla, Mayaguez, and Ponce MSAs)
- Non-MSA

The two Return Rate categories are:

- High
- Low

Return rate is a tract-level variable measuring the proportion of occupied housing units in the mailback universe which returned a census questionnaire. Low (high) return rate tracts are those tracts whose return rate is less than or equal to (greater than) the 25^{th} percentile return rate. Separate 25^{th} percentile cut-off values will be formed for the six applicable Race/Hispanic Origin by Tenure groups. Persons in List/Enumerate, Rural Update/Enumerate, and Urban Update/Enumerate TEAs are automatically placed in the High category. For Puerto Rico, distinct 25^{th} percentile return rate cut-off values will be formed for each Tenure category.

The four Region categories are:

- Northeast
- Midwest
- South
- West

B. Pre-collapsing

All Race/Hispanic Origin, Age/Sex, and Tenure categories for the U.S. will initially be maintained. The pre-collapsing plan for Region, MSA/TEA and Return Rate varies as follows:

- Non-Hispanic White or "Some other race" Owners: No collapsing
- Non-Hispanic White or "Some other race" Non-owners: Eliminate Region
- Non-Hispanic Black: Eliminate Region and partial collapsing of the MSA/TEA variable within Return Rate and Tenure categories
- Hispanic: Eliminate Region and partial collapsing of the MSA/TEA variable within Return Rate and Tenure categories
- Native Hawaiian or Pacific Islander: Eliminate the Region, Return Rate and MSA/TEA variables (Retain Tenure and Age/Sex only)
- Non-Hispanic Asian: Eliminate the Region, Return Rate and MSA/TEA variables (Retain Tenure and Age/Sex only)
- American Indian or Alaska Native on Reservations: Eliminate the Region, Return Rate and MSA/TEA variables (Retain Tenure and Age/Sex only)
- Off-Reservation American Indian or Alaska Native: Eliminate the Region, Return Rate and MSA/TEA variables (Retain Tenure and Age/Sex only)

For Puerto Rico, all 84 post-strata defined by MSA, Tenure, Return Rate, and Age/Sex will initially be maintained. Thus, there will be no pre-collapsing for Puerto Rico.

C. Post-collapsing

Depending on the actual A.C.E. sample sizes, additional collapsing may be necessary. The collapsing procedure is hierarchical which requires a pre-defined collapsing order. Given the pre-collapsing plan which yielded 448 post-strata, not much post-collapsing is anticipated. However, an extensive post-collapsing strategy is presented for completeness and to satisfy the requirement of pre-specification.

Note that collapsing does not necessarily imply elimination of a variable. Collapsing can refer to a reduction in the number of categories for a variable. For both the U.S. and Puerto Rico, a post-stratum is deemed too small if it contains fewer than 100 P Sample persons. The following general outline describes the post-collapsing hierarchy which is applied to both the U.S. and Puerto Rico. Any differences in definitions for Puerto Rico are noted.

If any of the 448 U.S. or 84 Puerto Rico post-strata are too small, Age/Sex will be collapsed first. This means that within any of the 64 U.S. (or 12 Puerto Rico) post-stratum groups, the seven Age/Sex categories defined in Table 1c will be reduced to the following three categories: Under 18, 18+ Male, and 18+ Female.

If some post-strata are still too small and require collapsing, Region will be collapsed next, if applicable. This collapsing applies only to the Non-Hispanic White or "Some other race" domain since the variable Region is only included in their post-stratification definition. In this case, all levels of Region (Northeast, Midwest, South, West) will be combined to eliminate the variable.

Next, the four-level MSA/TEA variable in the U.S. will be collapsed, if necessary, into the following two groups:

- Large and Medium MSA MO/MB
- Small MSA and Non-MSA MO/MB and All Other TEAs

For Puerto Rico, the three-level MSA variable will be collapsed, if necessary, into the following two groups:

- San Juan CMSA
- Other MSA and Non-MSA

If further collapsing is necessary, Return Rate is the next variable to collapse. High and Low Return Rate categories are combined to eliminate the variable.

Further collapsing involves the variable MSA/TEA in the U.S. (MSA in Puerto Rico). If necessary, the two groups defined above would be combined together to eliminate the variable MSA/TEA for the U.S. (MSA in Puerto Rico) completely.

The next variable to collapse is Tenure. Owner and Non-owner categories are combined to eliminate the variable entirely, if necessary.

If collapsing is still needed, the three remaining Age/Sex post-strata will be combined together to eliminate the Age/Sex variable completely.

In the event that there are not at least 100 P Sample persons in a Race/Hispanic Origin domain, all persons in that domain will be combined with Domain 7, which includes Non-Hispanic White and "Some other race."

V. RACE AND HISPANIC ORIGIN CLASSIFICATIONS

The Census 2000 questionnaire has 15 possible race responses. The 15 responses are collapsed into six major race groups as shown below. Races which are included in the major groups are shown in parentheses. Persons self-identifying with a single race essentially place themselves into one of these six categories.

- White
- Black (Black, African American, Negro)
- American Indian or Alaska Native
- Asian (Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Other Asian)
- Native Hawaiian or Pacific Islander (Native Hawaiian, Guamanian or Chamorro, Samoan, Other Pacific Islander)
- "Some other race"

For the first time in census history, persons will be able to respond to more than one race category. Allowing persons to self-identify with multiple races results in many more than six race groups. In fact, after collapsing race to the six major groups, there are $2^6 - 1 = 63$ possible race combinations. It is necessary to subtract the 1 in this equation since each individual is assumed to have a race.

The race variable defined above is often cross-classified with the Hispanic origin variable to define post-strata. The Hispanic origin variable consists of two responses, No and Yes. Categories which are included in the Yes response are shown in parentheses.

- No, not Spanish/Hispanic/Latino
- Yes (Mexican, Mexican American, Chicano, Puerto Rican, Cuban, Other Spanish/Hispanic/Latino)

Combining the race and Hispanic origin variables yields $63 \times 2 = 126$ possible Race/Hispanic Origin groups. It is important to note that any post-stratification plan of interest cannot support 126 Race/Hispanic Origin groups. As a solution, each of the 126 Race/Hispanic Origin response possibilities are assigned to one of seven Race/Hispanic Origin domains. The seven Race/Hispanic Origin domains are defined as follows:

- American Indian or Alaska Native on Reservations
- Off-Reservation American Indian or Alaska Native
- Hispanic
- Non-Hispanic Black
- Native Hawaiian or Pacific Islander
- Non-Hispanic Asian
- Non-Hispanic White or "Some other race"

Note that missing race and Hispanic origin data are imputed. Rules for classifying the 126 race and Hispanic origin combinations into one of the seven Race/Hispanic Origin domains are now presented.

Many of the decisions on how to classify multiple race persons are based on cultural, linguistic, and sociological factors which are known to affect coverage and are not necessarily data-driven.

A hierarchy is used to assign persons to a Race/Hispanic Origin domain. The Race/Hispanic Origin designation occurs in the following order: American Indian or Alaska Native on Reservations, Off-Reservation American Indian or Alaska Native, Hispanic, Non-Hispanic Black, Native Hawaiian or Pacific Islander, Non-Hispanic Asian, and Non-Hispanic White or "Some other race." All census data are tabulated using the race and Hispanic origin categories selected by census respondents.

For the following tables, Indian Country (IC) is a block-level variable that indicates whether a collection block is (wholly/partially) inside an American Indian reservation/trust land, Tribal Jurisdiction Statistical Area (TJSA), Tribal Designated Statistical Area (TDSA), or Alaska Native Village Statistical Area (ANVSA).

Tables 2 and 3 display the assignment of Race/Hispanic Origin domains. Table 2 applies to Hispanic persons while Table 3 applies to non-Hispanic persons. The first six rows of Tables 2 and 3 correspond to a single race response. The remaining portion of the tables addresses the assignment of multiple race responses to a single Race/Hispanic Origin domain. Although a person may be associated with multiple race responses, each person is included in only one of the seven Race/Hispanic Origin domains. All persons with a common number are assigned to the same Race/Hispanic Origin domain. Following is a verbal description of who is included in each Race/Hispanic Origin domain and their associated domain number.

Domain 1 (Includes American Indian or Alaska Native on Reservations): This domain includes any person living on a reservation marking American Indian or Alaska Native either as their single race or as one of many races, regardless of their Hispanic origin.

Domain 2 (Includes Off-Reservation American Indian or Alaska Native): This domain includes any person living in IC but not on a reservation who marks American Indian or Alaska Native either as their single race or as one of many races, regardless of their Hispanic origin. This domain also includes any non-Hispanic person not living in IC who marks American Indian or Alaska Native as their single race.

Domain 3 (Includes Hispanic): This domain includes all Hispanic persons who are not included in Domains 1 or 2. All Hispanic persons who self-identify with three or more races (excluding American Indian or Alaska Native in IC) are included in Domain 3. The only exception to this rule occurs when a Hispanic person lives in the state of Hawaii and classifies themselves as Native Hawaiian or Pacific Islander, regardless of whether they identify with a single or multiple race. All Hispanic persons satisfying this condition are re-classified into Domain 5.

Domain 4 (Includes Non-Hispanic Black): This domain includes any non-Hispanic person who marks Black as their only race. It also includes the combination of Black and American Indian or Alaska Native not in IC. In addition, people who mark Black and another single race group (Native Hawaiian or Pacific Islander, Asian, White, or "Some other race") are included in Domain 4. The only exception to this rule occurs when a Non-Hispanic Black person lives in the state of Hawaii and classifies themselves

as Native Hawaiian or Pacific Islander. All Non-Hispanic Black persons satisfying this condition are re-classified into Domain 5.

Domain 5 (Includes Native Hawaiian or Pacific Islander): This domain includes any person marking the single race Native Hawaiian or Pacific Islander. It also includes the combination of Native Hawaiian or Pacific Islander and American Indian or Alaska Native not in IC. Also included is the combination of Native Hawaiian or Pacific Islander with Asian. All persons living in the state of Hawaii who classify themselves as Native Hawaiian or Pacific Islander, regardless of their Hispanic origin and whether they identify with a single or multiple race, are also included in Domain 5.

Domain 6 (Includes Non-Hispanic Asian): This domain includes any non-Hispanic person marking Asian as their single race. If a person self-identifies with Asian and American Indian or Alaska Native not in IC, they are included in Domain 6.

Domain 7 (Includes Non-Hispanic White or "Some other race"): Non-Hispanic White or Non-Hispanic "Some other race" persons are included Domain 7. Non-Hispanic persons who self-identify with American Indian or Alaska Native not in IC and are White or "Some other race" are classified into Domain 7. If a Native Hawaiian or Pacific Islander response is combined with a White or "Some other race" response, they also are included in Domain 7. A person who self-identifies with Asian and White or Asian and "Some other race" is also included in this domain.
Finally, all non-Hispanic persons who self-identify with three or more races (excluding American Indian or Alaska Native in IC) are included in Domain 7. The only exception to this rule occurs when a Non-Hispanic White or Non-Hispanic "Some other race" person lives in Hawaii and classifies themselves as Native Hawaiian or Pacific Islander, regardless of whether they identify with other races. Persons who satisfy this criteria are re-classified into Domain 5.

Table 2: Census 2000 A.C.E. Post-stratification Domains for Hispanic

			Indian Country (IC)	
		Not in IC	Not On Res.	On Res.
Single race:				
American Indian or Alaska Native		3	2	1
Black		3	3	3
Native Hawaiian or Pacific Islander		3*	3	3
Asian		3	3	3
White		3	3	3
"Some other race"		3	3	3
American Indian or Alaska Native and:	Black	3	2	1
	Native Hawaiian or Pacific Islander	3*	2	1
	Asian	3	2	1
	White	3	2	1
	"Some other race"	3	2	1
Black and:	Native Hawaiian or Pacific Islander	3*	3	3
	Asian	3	3	3
	White	3	3	3
	"Some other race"	3	3	3
Native Hawaiian or Pacific Islander and:	Asian	3*	3	3
	White	3*	3	3
	"Some other race"	3*	3	3
Asian and:	White	3	3	3
	"Some other race"	3	3	3
American Indian or Alaska Native and:	Two or More Races	3*	2	1
All Else**		3*	3	3

* All persons living in the state of Hawaii who classify themselves as Native Hawaiian or Pacific Islander, regardless of their Hispanic origin and whether they identify with a single or multiple race, are included in Domain 5, which includes Native Hawaiian or Pacific Islander.

** All Else encompasses all remaining combinations which exclude American Indian or Alaska Native.

Table 3: Census 2000 A.C.E. Post-stratification Domains for Non-Hispanic

			Indian Country (IC)	
		Not in IC	Not On Res.	On Res.
Single race:				
American Indian or Alaska Native		2	2	1
Black		4	4	4
Native Hawaiian or Pacific Islander		5	5	5
Asian		6	6	6
White		7	7	7
"Some other race"		7	7	7
American Indian or Alaska Native and:	Black	4	2	1
	Native Hawaiian or Pacific Islander	5	2	1
	Asian	6	2	1
	White	7	2	1
	"Some other race"	7	2	1
Black and:	Native Hawaiian or Pacific Islander	4*	4	4
	Asian	4	4	4
	White	4	4	4
	"Some other race"	4	4	4
Native Hawaiian or Pacific Islander and:	Asian	5	5	5
	White	7*	7	7
	"Some other race"	7*	7	7
Asian and:	White	7	7	7
	"Some other race"	7	7	7
American Indian or Alaska Native and:	Two or More Races	7*	2	1
All Else**		7*	7	7

* All persons living in the state of Hawaii who classify themselves as Native Hawaiian or Pacific Islander, regardless of their Hispanic origin and whether they identify with a single or multiple race, are included in Domain 5, which includes Native Hawaiian or Pacific Islander.

** All Else encompasses all remaining combinations which exclude American Indian or Alaska Native.

ATTACHMENT: 1990 PES Post-Stratification

This attachment provides a brief summary of the 1990 PES post-stratification for the U.S. and Puerto Rico. Included below are the 51 post-stratum groups for the U.S. and the three post-stratum groups for Puerto Rico. Each of these post-stratum groups are further subdivided into the same seven age/sex groups.

Table 4a: 1990 PES 357 Design - 51 Post-Stratum Groups (U.S.)

Race/Hispanic Origin	Tenure	Urbanicity	N	M	S	W
Non-Hispanic White or Other	Owner	Large Urbanized Areas	1	2	3	4
		Other Urban	5	6	7	8
		Non-Urban	9	10	11	12
	Non-owner	Large Urbanized Areas	13	14	15	16
		Other Urban	17	18	19	20
		Non-Urban	21	22	23	24
Black	Owner	Large Urbanized Areas	25	26	27	28
		Other Urban	29			
		Non-Urban	30			
	Non-owner	Large Urbanized Areas	31	32	33	34
		Other Urban	35			
		Non-Urban	36			
Hispanic White or Other	Owner	Large Urbanized Areas	37	38	39	40
		Other Urban	41			
		Non-Urban	42			
	Non-owner	Large Urbanized Areas	43	44	45	46
		Other Urban	47			
		Non-Urban	48			
Asian or Pacific Islander	Owner		49			
	Non-owner		50			
Reservation Indians			51			

Table 4b: 1990 PES - 3 Post-Stratum Groups (Puerto Rico)

Place Type	
Central City in an MSA/PMSA	1
Non-central City in an MSA/PMSA	2
Not in an MSA/PMSA	3

Table 4c: 1990 PES - 7 Age/Sex Groups (U.S. & Puerto Rico)

	Male	Female
Under 18	A	
18 to 29	B	C
30 to 49	D	E
50+	F	G

Key:

MSA: Metropolitan Statistical Area, as defined by the Office of Management and Budget (OMB), will be referred to as MSA.

PMSA: Primary Metropolitan Statistical Area, as defined by the Office of Management and Budget (OMB), will be referred to as PMSA.

ESCAP MEETING NO. 7 - 04/12/00

MINUTES

ESCAP MEETING NO. 7 - 04/12/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 7

April 12, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The seventh meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on April 12, 2000 at 10:30. The agenda for the meeting was A.C.E. Post-Stratification.

Persons in attendance:

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
John Thompson
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura
Tommy Wright
Raj Singh
Gregg Robinson
Dawn Haines
Maria Urrutia
Annette Quinlan

I. Overview of A.C.E. Post-Stratification

Raj Singh and Dawn Haines discussed and distributed a draft memorandum describing the final post-stratification plan for A.C.E. Dual System Estimation. The memorandum details the definitions of the post-strata variables that will be used in the 2000 design.

Howard Hogan summarized how he had incorporated comments from a previous ESCAP

meeting on post-stratification into this memo, which has been finalized and is attached.

Howard indicated the main unresolved issue is how to collapse post-strata in the event of small A.C.E. sample sizes. This decision has not been finalized but the basic methodology will be completed by late April and the final specifications will be completed by June. All post-stratification issues and decisions will be made before the appropriate data are available.

The US post-stratification design for 2000 A.C.E. will contain a maximum of 448 post-strata, as compared to 357 post-strata for the 1990 post-stratification design. The major differences from the 1990 design are as follows: (1) the 2000 A.C.E. design includes the new variable mail return rate, (2) region is included only for Non-Hispanic White or "Some Other Race" Owners, (3) the 1990 PES urbanicity variable has been redefined by combining Metropolitan Statistical Area (MSA) by Type of Enumeration Area (TEA), (4) TEA is also a new variable for 2000 A.C.E., and (5) the document describes in detail the treatment of multiple race responses for creating post-stratification domains.

The Puerto Rico A.C.E. post-stratification plan has a maximum of 84 post-strata. The variables Tenure and Return Rate have both been added since 1990. A post-collapsing plan for Puerto Rico will also be defined before data are available.

II. Next Meeting

The next meeting scheduled for Wednesday April 26, 2000 will discuss A.C.E. Weight Trimming.

ESCAP Committee

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogen
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 8 - 04/26/00

AGENDA

Kathleen P Zveare
04/25/2000 03:21 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC

cc:

Subject: Agenda for 4/26 ESCAP Meeting

The agenda for tomorrow's ESCAP meeting is:

Weight Trimming - Donna Kostanich

Time: 10:30-12:00

Room: 2412/3

ESCAP MEETING NO. 8 - 04/26/00

HANDOUTS

Thoughts on A.C.E. Weight Trimming — DRAFT 4/26/00

Goal: Reduce the contribution of variance due to outlier clusters dominating a post-stratum's coverage correction factor.

Key Dates:

June, 2000 Determine final weight trimming plans.

Dec. 8, 2000 Apply and verify the weights to Missing Data Files.

Assumptions:

- **S** Cluster level trimming
- **S** Identify clusters to downweight and proportionately upweight the remaining clusters.
- **S** Identify outlier clusters separately for American Indian Reservations, rest of the U.S., and Puerto Rico.
- **S** Distribute weights within sampling stratum if feasible.
- **S** Sampling Staff will write the programs to identify clusters and do the weight trimming. This requires running parallel systems to verify the results. The weights will be transmitted to programmers electronically to be applied to Missing Data files.

Three options:

1. Trim weights based on total weighted housing unit estimates of block clusters
   - implement trimming methodology: Oct. 21 to Nov. 30
   - separately for P & E samples
   - does not reflect any matching results
   - does not reflect TES

2. Trim weights using the initial housing unit match results
   - implement trimming methodology: Oct. 21 to Nov. 30
   - use housing unit match results as proxy for person matching
   - will not reflect changes to census since the January DMAF
   - could misidentify clusters to downweight
   - does not reflect TES, but could if willing to make a guess at effect of TES

3. Trim weights using the person match results
   - implement trimming methodology: Dec. 1 to Dec. 5
   - reflects impact of TES
   - relies on person matching ending on Nov. 30
   - can use HU match results as a contingency

April 26, 2000 Draft

Potentially Influential Cluster

TEA:	Urban Update/Leave
1990 Housing Unit Count:	217
Keyed and Valid Independent Listing Count:	192
January DMAF	1153

Housing Unit Matching Results

Matches	192
Erroneous Enumerations	961

1990 Demographic/Tenure Distributions

Black Renters	97.0%
Hispanic Renters	1.7%
Black Owners	0.8%
Other Renters	0.5%

Final P-sample Weight:	383.130

ESCAP MEETING NO. 8 - 04/26/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 8

April 26, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The eighth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on April 26, 2000 at 10:30. The agenda for the meeting was A.C.E. Weight Trimming.

Persons in attendance:

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan
Susan Miskura
Donna Kostanich
Raj Singh
Gregg Robinson
Signe Wetrogan
Carolee Bush
Maria Urrutia
Annette Quinlan

I. A.C.E. Weight Trimming

Donna Kostanich presented options for A.C.E. weight trimming. The major goal of weight trimming would be to reduce the effect on the A.C.E. estimates due to outlier clusters dominating a post-stratum's coverage correction factor. The assumptions and requirements for weight trimming were discussed, as were alternatives. To facilitate understanding the potential effects of weight trimming, an example of an influential cluster was distributed and discussed. The example and options for weight trimming are attached.

The ESCAP discussed three options for implementing weight trimming:

1) Identify block clusters for weight trimming based on total weighted housing unit estimates of block clusters from the initial housing unit phase. For example, a cluster with a large weighted estimate of housing units from the initial phase would be identified for weight trimming.
2) Identify block clusters for weight trimming using the results from the initial housing unit match that occurred in April 2000. For example, these block clusters would be identified if they had a large number of non-matching housing units.
3) Identify block clusters for weight trimming based on the results of the person match that will be completed in November 2000. For example, these block clusters would be identified if they had a large number of non-matched person records.

The timing for alternatives 1 and 2 would occur from 10/21/00 to 11/30/00 since they do not require the person match results. Since these alternatives do not include the interviewing results they would not identify all clusters that may require weight trimming. Given that, we decided that alternative 3 would be the best option to use for weight trimming. We noted, however, that there were timing concerns because this option will occur from 12/1/00 to 12/5/00, allowing five days for implementation. Therefore, there is a risk of extending the A.C.E. schedule if this process requires more than five days to complete, including the necessary review process.

It was also decided that the weight trimming process, including the criteria, would be identified by June. It is critical that this process and criteria be pre-specified and publicly available. Before any weight trimming would be implemented, the ESCAP will review to ensure that pre-specified criteria are met.

II. Next Meeting

The next meeting scheduled for Wednesday May 24, 2000 will discuss Telephone Interviewing and Synthetic Estimation.

ESCAP Committee

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogen
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 9 - 05/24/00

AGENDA

Kathleen P Zveare
05/23/2000 02:32 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC
cc:
Subject: Agenda for 5/24 ESCAP Meeting

The agenda for the May 24 ESCAP meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

1. Telephone Interviewing
2. Synthetic Estimation

ESCAP MEETING NO. 9 - 05/24/00

HANDOUTS

OVERVIEW OF SYNTHETIC ESTIMATION

Dawn Haines

May 22, 2000

Goal: Obtain an integer number of persons for each post-stratum within each tabulation block, representing either overcounts or undercounts.

- Synthetic estimation begins after Dual System Estimates are calculated for each post-stratum.

- The procedure involves carrying down and rounding from post-strata to tabulation blocks.

- This occurs at the following levels:

  - State
  - County
  - Tract
  - Block

- Finally, the number of records to replicate for each post-stratum within each tabulation block is determined.

May 22, 2000 **CONTROLLED ROUNDING EXAMPLE**

Dawn Haines

Suppose our nation is made up of the following i = 5 post-strata: White, Black, AIAN, API, and Other. Dual System Estimates (DSE_i), census counts (C_i), and coverage correction factors (CCF_i) are given for each post-stratum i. DSE and census totals over post-strata are presented.

Table 1: Post-stratum Information

i	White	Black	AIAN	API	Other	Total
DSE_i	3,733,740.22	433,730.80	102,864.39	80,277.83	115,871.70	4,466,484.94
C_i	3,809,939	409,180	94,371	79,483	110,354	4,503,327
CCF_i	0.98	1.06	1.09	1.01	1.05	

The coverage correction factor for post-stratum i, CCF_i, is formed by dividing the DSE for post-stratum i by its census count, denoted

$$CCF_i = \frac{DSE_i}{C_i}.$$

Table 2: Control-rounded Dual System Estimates

	White	Black	AIAN	API	Other	Total
DSE_i^R	3,733,740	433,731	102,864	80,278	115,872	4,466,485

Each Level of Carrying down requires 4 steps:

1) Census Counts: $C_{i,s}$

2) Synthetic: $\hat{N}_{i,s}^{S} = C_{i,s} \times CCF_i$

3) Adjusted Synthetic: $\hat{N}_{i,s}^{AS} = \hat{N}_{i,s}^{S} \times \frac{DSE_i^R}{DSE_i}$

(Note that the ratio of the rounded to the unrounded DSE is replaced by the ratio of rounded to unrounded Synthetic for all levels of carrying down except for the 1st level State.)

4) Rounded Synthetic: $\hat{N}_{i,s}^{RS}$ (Control rounding of Adjusted Synthetic.)

3 States in U.S.

Table 3: State-level Census Counts $C_{i,s}$

	White	Black	AIAN	API	Other	Total
02 (AK)	415,492	22,451	85,698	19,728	6,674	550,043
09 (CT)	2,859,353	274,269	6,654	50,698	96,142	3,287,116
10 (DE)	535,094	112,460	2,019	9,057	7,538	666,168
Total	3,809,939	409,180	94,371	79,483	110,354	4,503,327

Table 4: State-level Synthetic Estimates

$$\hat{N}^{S}_{i,s} = C_{i,s} \times CCF_i \quad \textbf{(9147.57 = 9057 × 1.01)}$$

	White	Black	AIAN	API	Other	Total
02 (AK)	407,182.16	23,798.06	93,410.82	19,925.28	7,007.70	551,324.02
09 (CT)	2,802,165.94	290,725.14	7,252.86	51,204.98	100,949.10	3,252,298.02
10 (DE)	524,392.12	119,207.60	2,200.71	**9,147.57**	7,914.90	662,862.90
Total	3,733,740.22	433,730.80	102,864.39	80,277.83	115,871.70	4,466,484.94

Table 5: State-level Adjusted Synthetic Estimates

$$\hat{N}^{AS}_{i,s} = \hat{N}^{S}_{i,s} \times \frac{DSE^{R}_{i}}{DSE_i} \quad \textbf{(9147.59 = 9147.57 × } \frac{80278}{80277.83} \textbf{)}$$

	White	Black	AIAN	API	Other	Total
02 (AK)	407,182.14	23,798.07	93,410.47	19,925.32	7,007.72	551,323.71
09 (CT)	2,802,165.78	290,725.27	7,252.83	51,205.09	100,949.36	3,252,298.33
10 (DE)	524,392.09	119,207.66	2,200.70	**9,147.59**	7,914.92	662,862.96
Total	3,733,740.00	433,731.00	102,864.00	80,278.00	115,872.00	4,466,485.00

Table 6: State-level Control-rounded Synthetic Estimates $\hat{N}_{i,s}^{RS}$

	White	Black	AIAN	API	Other	Total
02 (AK)	407,182	23,798	93,411	19,925	7,008	551,324
09 (CT)	2,802,166	290,725	7,253	51,205	100,949	3,252,298
10 (DE)	524,392	119,208	2,200	9,148	7,915	662,863
Total	3,733,740	433,731	102,864	80,278	115,872	4,466,485

3 Counties in Delaware

Table 7: County-level Census Counts $C_{i,c}$

DE	White	Black	AIAN	API	Other	Total
001	87,300	20,631	614	**1,420**	1,028	110,993
003	355,399	72,834	760	7,048	5,905	441,946
005	92,395	18,995	645	589	605	113,229
Total	535,094	112,460	2,019	9,057	7,538	666,168

Table 8: County-level Synthetic Estimates

$\hat{N}_{i,c}^{S} = C_{i,c} \times CCF_i$ **(1434.20 = 1420 × 1.01)**

DE	White	Black	AIAN	API	Other	Total
001	85,554.00	21,868.86	669.26	**1,434.20**	1,079.40	110,605.72
003	348,291.02	77,204.04	828.40	7,118.48	6,200.25	439,642.19
005	90,547.10	20,134.70	703.05	594.89	635.25	112,614.99
Total	524,392.12	119,207.60	2,200.71	**9,147.57**	7,914.90	662,862.90

Table 9: County-level Adjusted Synthetic Estimates

$$\hat{N}^{AS}_{i,c} = \hat{N}^{S}_{i,c} \times \frac{\hat{N}^{RS}_{i,s}}{\hat{N}^{S}_{i,s}} \quad (\mathbf{1434.27 = 1434.20} \times \frac{9148}{9147.57})$$

DE	White	Black	AIAN	API	Other	Total
001	85,553.98	21,868.93	669.04	**1,434.27**	1,079.41	110,605.64
003	348,290.94	77,204.30	828.13	7,118.81	6,200.33	439,642.52
005	90,547.08	20,134.77	702.82	594.92	635.26	112,614.85
Total	524,392.00	119,208.00	2,200.00	**9,148.00**	7,915.00	662,863.00

Table 10: County-level Control-rounded Adjusted Synthetic Estimates $\hat{N}^{RS}_{i,c}$

DE	White	Black	AIAN	API	Other	Total
001	85,554	21,869	669	**1,434**	1,080	110,606
003	348,291	77,204	828	7,119	6,200	439,642
005	90,547	20,135	703	595	635	112,615
Total	524,392	119,208	2,200	**9,148**	7,915	662,863

4 Tracts in County 1 in Delaware

Table 11: Tract-level Census Counts $C_{i,t}$

Co. 1, DE	White	Black	AIAN	API	Other	Total
tract 1	29,004	6,854	216	422	239	36,735
tract 2	5,408	9,315	173	**314**	411	15,621
tract 3	37,816	3,298	175	477	113	41,879
tract 4	15,072	1,164	50	207	265	16,758
Total	87,300	20,631	614	1,420	1,028	110,993

Table 12: Tract-level Synthetic Estimates

$$\hat{N}^{S}_{i,t} = C_{i,t} \times CCF_i \quad (\mathbf{317.14 = 314 \times 1.01})$$

Co. 1, DE	White	Black	AIAN	API	Other	Total
tract 1	28,423.92	7,265.24	235.44	426.22	250.95	36,601.77
tract 2	5,299.84	9,873.90	188.57	**317.14**	431.55	16,111.00
tract 3	37,059.68	3,495.88	190.75	481.77	118.65	41,346.73
tract 4	14,770.56	1,233.84	54.50	209.07	278.25	16,546.22
Total	85,554.00	21,868.86	669.26	**1,434.20**	1,079.40	110,605.72

Table 13: Tract-level Adjusted Synthetic Estimates

$$\hat{N}^{AS}_{i,t} = \hat{N}^{S}_{i,t} \times \frac{\hat{N}^{RS}_{i,c}}{\hat{N}^{S}_{i,c}} \quad \left(\mathbf{317.10 = 317.14 \times \frac{1434}{1434.20}}\right)$$

Co. 1, DE	White	Black	AIAN	API	Other	Total
tract 1	28,423.92	7,265.29	235.35	426.16	251.09	36,601.81
tract 2	5,299.84	9,873.96	188.50	**317.10**	431.79	16,111.19
tract 3	37,059.68	3,495.90	190.68	481.70	118.72	41,346.68
tract 4	14,770.56	1,233.85	54.48	209.04	278.40	16,546.33
Total	85,554.00	21,869.00	669.00	**1,434.00**	1,080.00	110,606.00

Table 14: Tract-level Control-rounded Adjusted Synthetic Estimates $\hat{N}^{RS}_{i,t}$

Co. 1, DE	White	Black	AIAN	API	Other	Total
tract 1	28,424	7,265	236	426	251	36,602
tract 2	5,300	9,874	188	317	432	16,111
tract 3	37,059	3,496	191	482	119	41,347
tract 4	14,771	1,234	54	209	278	16,546
Total	85,554	21,869	669	1,434	1,080	110,606

3 Blocks in Tract 2 in County 1 in Delaware

Table 15: Block-level Census Counts $C_{i,b}$

tract 2, Co. 1,DE	White	Black	AIAN	API	Other	Total
block 1	1,785	2,508	32	150	128	4,603
block 2	1,429	4,283	71	64	245	6,092
block 3	2,194	2,524	70	**100**	38	4,926
Total	5,408	9,315	173	314	411	15,621

Table 16: Block-level Synthetic Estimates

$\hat{N}^{S}_{i,b} = C_{i,b} \times CCF_i$ ***(101.00 = 100 × 1.01)***

tract 2 Co. 1, DE	White	Black	AIAN	API	Other	Total
block 1	1,749.30	2,658.48	34.88	151.50	134.40	4,728.56
block 2	1,400.42	4,539.98	77.39	64.64	257.25	6,339.68
block 3	2,150.12	2,675.44	76.30	***101.00***	39.90	5,042.76
Total	5,299.84	9,873.90	188.57	317.14	431.55	16,111.00

Table 17: Block-level Adjusted Synthetic Estimates

$\hat{N}^{AS}_{i,b} = \hat{N}^{S}_{i,b} \times \frac{\hat{N}^{RS}_{i,t}}{\hat{N}^{S}_{i,t}}$ $\left(100.96 = 101.00 \times \frac{317}{317.14} \right)$

tract 2 Co. 1, DE	White	Black	AIAN	API	Other	Total
block 1	1,749.35	2,658.51	34.77	151.43	134.54	4,728.61
block 2	1,400.46	4,540.03	77.16	64.61	257.52	6,339.77
block 3	2,150.18	2,675.47	76.07	***100.96***	39.94	5,042.62
Total	5,300.00	9,874.00	188.00	317.00	432.00	16,111.00

Table 18: Block-level Control-rounded Adjusted Synthetic Estimates $\hat{N}_{i,b}^{RS}$

tract 2 Co. 1, DE	White	Black	AIAN	API	Other	Total
block 1	1,749	2,659	35	151	135	4,729
block 2	1,401	4,540	77	65	257	6,340
block 3	2,150	2,675	76	101	40	5,042
Total	5,300	9,874	188	317	432	16,111

Number of Records to Create

Table 19: Difference Between Block-level Control-rounded Adjusted Synthetic Estimates and Census Counts

$$\hat{N}_{i,b}^{RS} - C_{i,b}$$

tract 2 in Co. 1, DE	White	Black	AIAN	API	Other	Total
block 1	- 36	151	3	1	7	126
block 2	- 28	257	6	1	12	248
block 3	- 44	151	6	1	2	116
Total	- 108	559	15	3	21	490

ESCAP MEETING NO. 9 - 05/24/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 9

May 24, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The ninth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on May 24, 2000 at 10:30. The agenda for the meeting was A.C.E. Telephone Interviewing and A.C.E. Synthetic Estimation.

Persons in attendance:

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura
Donna Kostanich
Raj Singh
Tommy Wright
Carolee Bush
Maria Urrutia
Annette Quinlan

I. A.C.E. Telephone Interviewing

Howard Hogan discussed the goals and concerns of A.C.E. telephone interviewing. Cases are selected for the telephone phase only if the A.C.E. housing unit was independently listed in the A.C.E., was matched to the census during the initial phase of the A.C.E. Housing Unit Matching operation, and returned their Census 2000 questionnaire with a valid telephone number.

The benefits of conducting the telephone interviews include:

(1) Full operational test of the CAPI instrument and its control and support system earlier in the process.
(2) A reduced number of movers between Census Day and interview day.
(3) The interviews may be more accurate because they are occurring closer to Census Day.
(4) Training supervisory staff on the CAPI instrument to provide an opportunity for them to become more familiar with the instrument and its functions before the person visit interviews are conducted.

The potential concerns of conducting telephone interviews are:

(1) The possible mode effect on the completeness and the data quality of the interview. For example, is there a difference on the completeness of the interview when it is conducted over the phone as compared to a personal visit.
(2) The possible lack of independence between the census and the A.C.E. There is a possibility that the respondent can remember how they answered similar questions during NRFU of late mail returns or other later census operations.

These potential concerns are not anticipated to have an appreciable impact on the estimates. The mode effect will not impact the estimates because (1) the same instrument that is used for telephoning is also used for the personal visit interviews, (2) if there is any resistence, the interview goes to the field for those households which do not want to respond by telephone, and (3) the telephone interviews represent the more cooperative households. The independence concern will not impact the estimates because (1) we have already matched the A.C.E. address list to the Census and we only conduct the telephone interviews at matched households and (2) the telephone interviews are only conducted in areas where there is little risk of mail delivery problems. Staff will be evaluating the results of the A.C.E. telephone interview operation.

II. Synthetic Estimation

Donna Kostanich described synthetic, or indirect, estimation and provided an example, which is attached. The goal of synthetic estimation is to carry down coverage correction factors from the DSE to the block level. We use a control rounding procedure, as we do for long form estimation, to ensure that corrections are made in the form of an integer.

The statistical correction that results from the A.C.E. is carried down to census blocks by applying the coverage correction factors within each A.C.E. post-stratum. The goal in constructing post-strata is to form groupings of the population that capture differences in the probabilities of being included in the census and the A.C.E. In effect, the inclusion probabilities

are more similar for individuals within the same post-stratum than for individuals in different post-strata. The coverage correction factors are calculated for each post-stratum, based on a representative sample of the post-stratum, and thus reflect the net coverage of all people within the post-stratum. This is the underlying basis for applying this factor to the data records within the corresponding post-stratum to produce statistically corrected block totals which serve as the basis for Census 2000 tabulations.

The accuracy of the estimates that result from the application of the coverage correction factors depends on the degree to which the net coverage for areas or groups within a post-stratum is similar to the coverage correction factor that was developed for that post-stratum. The coverage correction factor is measured for the post-stratum based on a representative sample, and thus represents the net coverage for the post-stratum. Clearly, within the post-stratum, some degree of variation is expected from the measured coverage correction factor, and this variation will most likely be relatively greater for small areas. Thus, it is inevitable that the A.C.E. will result in the population in some blocks being overestimated and the population in other blocks being underestimated. The A.C.E. statistical correction was never intended nor expected to produce unqualified improvement in the smallest geographic areas, like blocks. That the A.C.E. does not produce improvement for every single block, however, is no reason to forego the benefits that will flow from the use of corrected census population counts at geographic levels of significance to data users. The Census Bureau expects that the A.C.E. estimates will produce better data for aggregations - such as states, congressional districts, counties, and cities - that are the basic areas for which census data are used.

The controlled rounding program, as described above, integerizes unrounded synthetic estimates. The rounding occurs in stages from (1) state, (2) county, (3) tract, and then by (4) block level. At each stage the controlled rounded estimates will differ by less than one from the unrounded estimator. The software is being double programmed to verify the results.

Bob Fay discussed the 1990 PES synthetic estimation assumptions, and his analysis of the effects of synthetic estimation on the analysis of the accuracy of the 1990 PES.

Application of synthetic estimation inevitably results in some degree of heterogeneity bias in estimates for states, counties, and other geographic areas of interest. The total error model of Mary Mulry and Bruce Spencer did not attempt to account for synthetic estimation bias, basically because the 1990 PES data were too thin to provide any reliable direct measures of this source of error.

At the time, there were competing hypotheses or possibilities:

(1) Heterogeneity, although omitted, might be so small as to be negligible.
(2) Heterogeneity might be large and, because it was omitted from the model, the error in the

adjusted figures might be much larger than estimated. Therefore, the loss function analysis might incorrectly favor adjustment.

Bob Fay attempted a brief summary of an empirical study presented in a 1993 paper coauthored by John Thompson. The analysis focused on the possible effect of heterogeneity on the loss function analysis at the state level. By taking census variables, such as mail response rates, poverty, and unemployment, and scaling them to the approximate level of the percent undercount, they constructed eight artificial populations. Unlike the PES, the artificial populations were essentially unaffected by sampling variance, since the variables were measured either by the whole census or the long form. In other words, heterogeneity bias could be studied for the artificial populations without the limitations imposed by the sampling error in the 1990 PES data. The loss function analysis, mimicking the 1990 PES by omitting any allowance for heterogeneity, could then be compared to the actual losses with and without adjustment, including the effect of heterogeneity in the calculation. In seven out of eight of the artificial populations, the results indicated the following:

> Heterogeneity, although a potentially significant source of error, led the loss function to understate the error in the unadjusted census counts by about as much as, or even more than, the error in the adjusted counts. Hence, although it omitted an important source of error, the loss function analysis generally could be trusted when it showed an advantage to the adjusted counts over the unadjusted. The one exception was the artificial population based on unemployment rate.

Bob noted that, resources permitting, it would be helpful to replicate this sort of study with 2000 data at some point.

III. Next Meeting

The next meeting scheduled for Wednesday June 28, 2000 will discuss Missing Data and Correlation Bias.

ESCAP Committee

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogen
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 10 - 06/28/00

AGENDA

Kathleen P Zveare
06/26/2000 01:14 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC
cc: Patrick J Cantwell/DSSD/HQ/BOC@BOC, J Gregory Robinson/POP/HQ/BOC@BOC
Subject: Agenda for 6/28 ESCAP Meeting

Below is the agenda for the June 28 ESCAP Meeting scheduled from 10:30-12 in Rm. 2412/3:

1. Missing Data---Pat Cantwell

2. Overview of Correlation Bias---Raj Singh

3. How Do We Measure Correlation Bias Using DA---Greg Robinson.

ESCAP MEETING NO. 10 - 06/28/00

HANDOUTS

Description of Before-Followup (BFU) Groups for the E-Sample

BFU Group 1. *Matches needing FU.* This group contains all E-Sample persons who match to a P-Sample person, but are sent to followup. This situation occurs when a person matches, but the P-Sample person to whom he or she matches has unresolved residence status.

BFU Group 2. *Possible matches.* This group contains all E-Sample persons with a match code of "possible match".

BFU Group 3. *Partial HH nonmatches.* This group contains all E-Sample persons unresolved before followup who did not match to a P-Sample person but who were living in a household where at least one person did match.

BFU Group 4. *Whole HH nonmatches (where HU matched); not conflicting HHs.* This group contains all E-Sample persons unresolved before followup who lived in a household where no persons matched but the housing unit did match to an A.C.E. housing unit. Note that this group does not contain persons from conflicting households*.

BFU Group 5. *Nonmatches from conflicting HHs; HU not in regular NRFU.* This group contains all E-Sample persons unresolved before followup who lived in a household where no person matched, the housing unit was **not** in regular nonresponse followup, and the household is a conflicting household*.

BFU Group 6. *Nonmatches from conflicting HHs; HU in regular NRFU.* This group contains all E-Sample persons unresolved before followup who lived in a household where no person matched, the housing unit **was** in regular nonresponse followup, and the household is a conflicting household*.

BFU Group 7. *Whole HH nonmatches; HU did not match during HU matching.* This group contains all E-Sample persons unresolved before followup who lived in a household where no person matched and the census housing unit did not match to an A.C.E. housing unit.

BFU Group 8. *Persons resolved before FU.* This group contains all E-Sample persons whose enumeration status was resolved before followup. The following people are included:

- **!** Matches not needing followup.
- **!** Duplicates of another E-Sample person
- **!** Persons erroneously enumerated due to geocoding error.
- **!** Persons identified as fictitious before followup.

BFU Group 9. *Persons with insufficient information for matching.* This group includes all persons who do not have a full name and at least 2 person characteristics.

*Conflicting households are those in which the census household has all different persons from the matching A.C.E. household. All persons in the census household and the A.C.E. household are nonmatches (no possible matches). Note that persons in both the census household and the A.C.E. household can have insufficient information for matching in a conflicting household.

Description of Before-Followup (BFU) Groups P-Sample

BFU Group 1. *Matches needing FU.* This group contains all P-Sample persons who match to an E-Sample person, but are sent to followup. This situation occurs when a person matches, but has unresolved residence status.

BFU Group 2. *Possible matches.* This group contains all P-Sample persons with a match code of "possible match".

BFU Group 3. *Partial HH nonmatches.* This group contains all P-Sample persons needing followup who did not match to an E-Sample person but who were living in a household where at least one person did match.

BFU Group 4. *Whole HH nonmatches needing FU (not conflicting HHs).* This group contains all P-Sample persons needing followup who lived in a household where no persons matched but the household is **not** conflicting*.

BFU Group 5. *Nonmatches from conflicting HHs needing FU.* This group contains all P-Sample persons needing followup who lived in a household where no person matched and the household is a conflicting* household.

BFU Group 6. *Persons resolved before FU.* This group contains all P-Sample persons whose residence status was resolved before followup. The following people are included:

- Matches not needing followup.
- Nonmatches where the data were collected from a household member (i.e., a nonproxy interview), the household was not conflicting*, and the residence status is known.
- Partial household nonmatches with a code of "NC".
- Duplicates of another P-Sample person.

BFU Group 7. *Persons with insufficient information for matching.* This group includes all persons who do not have a full name and at least 2 person characteristics.

*Conflicting households are those in which the census household has all different persons from the matching A.C.E. household. All persons in the census household and the A.C.E. household are nonmatches (no possible matches). Note that persons in both the census household and the A.C.E. household can have insufficient information for matching in a conflicting household.

Missing Data for the Accuracy and Coverage Evaluation

June 28, 2000

1. Background: Types of missing data

- Noninterviews (P Sample only)
  - --- census day
  - --- A.C.E. interview day

- Item nonresponse (P Sample only)
  only for tenure, race, hispanic origin, age category, sex

- Unresolved status
  - --- correct enumeration (E Sample)
  - --- match (P Sample)
  - --- residence (P Sample)

2. Noninterviews (housing-unit level)

compute two adjustments of sampling weights at housing-unit level (see Attachment 1):

- one for census day, applied to person non-movers and out-movers
- one for A.C.E. interview day, applied to person in-movers

3. Item nonresponse

characteristic imputation (see Attachment 2)

tenure, race, hispanic origin: nearest-neighbor hot deck

age, sex: impute from distribution of characteristic,
conditioned on certain variables

4. Unresolved status

- choice of imputation cells for the U.S. (see Attachment 3)
  criteria
  - --- variables that discriminate well
  - --- minimum expected frequencies of resolved cases per cell
  - --- no collapsing

  research
  - --- 1990 PES, Hudson and Clarke
  - --- 1998, Dress Rehearsal, Malec

- Puerto Rico (see Attachment 3)
  - --- start with the same cells
  - --- remove race, number of imputes
  - --- maintain expected frequencies

- procedures used in the 1990s
  - 1990 PES: logistic regression modeling; Belin, Diffendal, Fay
  - 1995, 1996, 1998: increasing use of imputation cell estimation
    - --- 1995, for residence status
    - --- 1996, for match status
    - --- 1998 (Dress Rehearsal), for all three statuses

- original possibilities for using logistic regression in the 2000 A.C.E.
  - --- Belin's program with coding changes
  - --- vendor software package
  - --- a new program written for the A.C.E.

- reasons for using imputation cell estimation in production for the A.C.E.
  - --- logistic regression options not acceptable
  - --- differences between logistic regression and imputation cell estimation appear to be small, based on 1990 PES, 1995 and 1996

5. Verification

a) verifying that the programs are running correctly
   - --- production program, all components (except for characteristic imputation) already double programmed
   - --- independent program being written

b) operational analyses
   - hundreds of computations and tables for review
   - statistical analyses done automatically, with unusual results raising flags

c) detailed analyses
   - additional tables also available

We define an interview and a noninterview. For the given reference date, that is, separately for census day and A.C.E. interview day,

> interview: a housing unit is an interview if there is at least one person (with name and at least two demographic characteristics) who possibly or definitely was a resident of the housing unit on the given reference date;
>
> noninterview: an occupied housing unit that is not an interview is a noninterview.

(An example on the next page illustrates these definitions and the procedure below.)

Procedure

1. Assign all occupied housing units to noninterview adjustment cells: block cluster × type of basic address (single-family home, apartment, or other). See note below on collapsing.

2. Assign *every* occupied unit a census-day interview status (interview or noninterview) *and* an A.C.E. interview-day interview status.

3. The noninterview adjustment factor is the ratio of the *weighted* number of housing-unit interviews and noninterviews to the weighted number of housing-unit interviews.

4. Compute the adjustment for census day; apply it to the person weights of non-movers and out-movers in interviewed housing units.

5. Compute the adjustment for A.C.E. interview day; apply it to the person weights of in-movers in interviewed housing units.

Notes:

- Although interview status is a housing-unit characteristic, mover status is a person characteristic. People in the same household can have different mover statuses.

- Vacant housing units do not contribute to the noninterview adjustment.

- People in noninterviewed housing units do not contribute to the components of DSE.

- If the *unweighted* number of interviewed housing units in a cell is less than half the number of noninterviewed units, the cell is collapsed. Rules for collapsing are in Q-25.

Suppose a block cluster has nine housing units, all of the same type of basic address, for example, all single family homes, as depicted below.

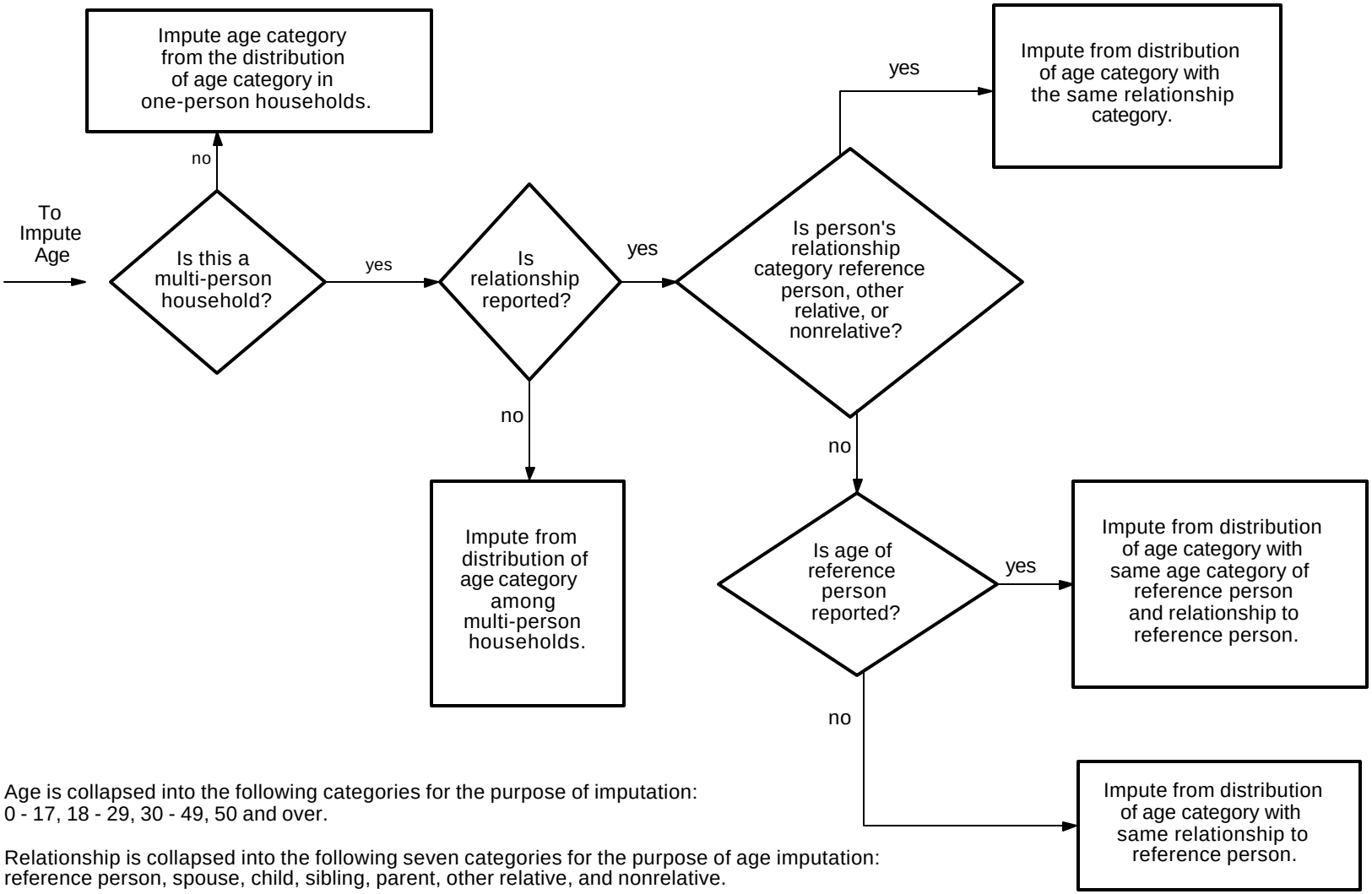
Housing Unit	Weight	Actual Situation	Status of (and Information from) A.C.E. Interview	Census-Day Status [1]	A.C.E. Interview-Day Status [1]
1	100	Resident on 4/1/00 and at time of A.C.E. intvw.	Interviewed in A.C.E.	I	I
2	100	Resident on 4/1 and at time of A.C.E. intvw.	Neighbor (proxy) interviewed in A.C.E.	I	I
3	100	Resident on 4/1 and at time of A.C.E. intvw.	No one interviewed in A.C.E.	NI	NI
4	100	Vacant on 4/1, resident at time of A.C.E. intvw.	Interviewed in A.C.E., knows of 4/1 status	Vacant	I
5	100	Vacant on 4/1, resident at time of A.C.E. intvw.	Intvw'd in A.C.E., no knowledge of 4/1 status	NI	I
6	100	Vacant on 4/1, resident at time of A.C.E. intvw.	No one interviewed in A.C.E.	NI	NI
7	100	Resident on 4/1, vacant at time of A.C.E. intvw.	Information obtained from proxy	I	Vacant
8	100	Resident on 4/1, vacant at time of A.C.E. intvw.	No info on 4/1 status; Census staff determines vacant at time of A.C.E.	NI	Vacant
9	100	Resident on 4/1, different resident at time of A.C.E.	Interviewed in A.C.E., knows of 4/1status	I	I

[1] Interview Status: I = interview, NI = noninterview

In this cluster × TBA (noninterview cell), to people in interviewed housing units, apply the following noninterview adjustments:

(1) to the person weights of non-movers and out-movers,
census-day NI adjustment = 800 / 400 = 2

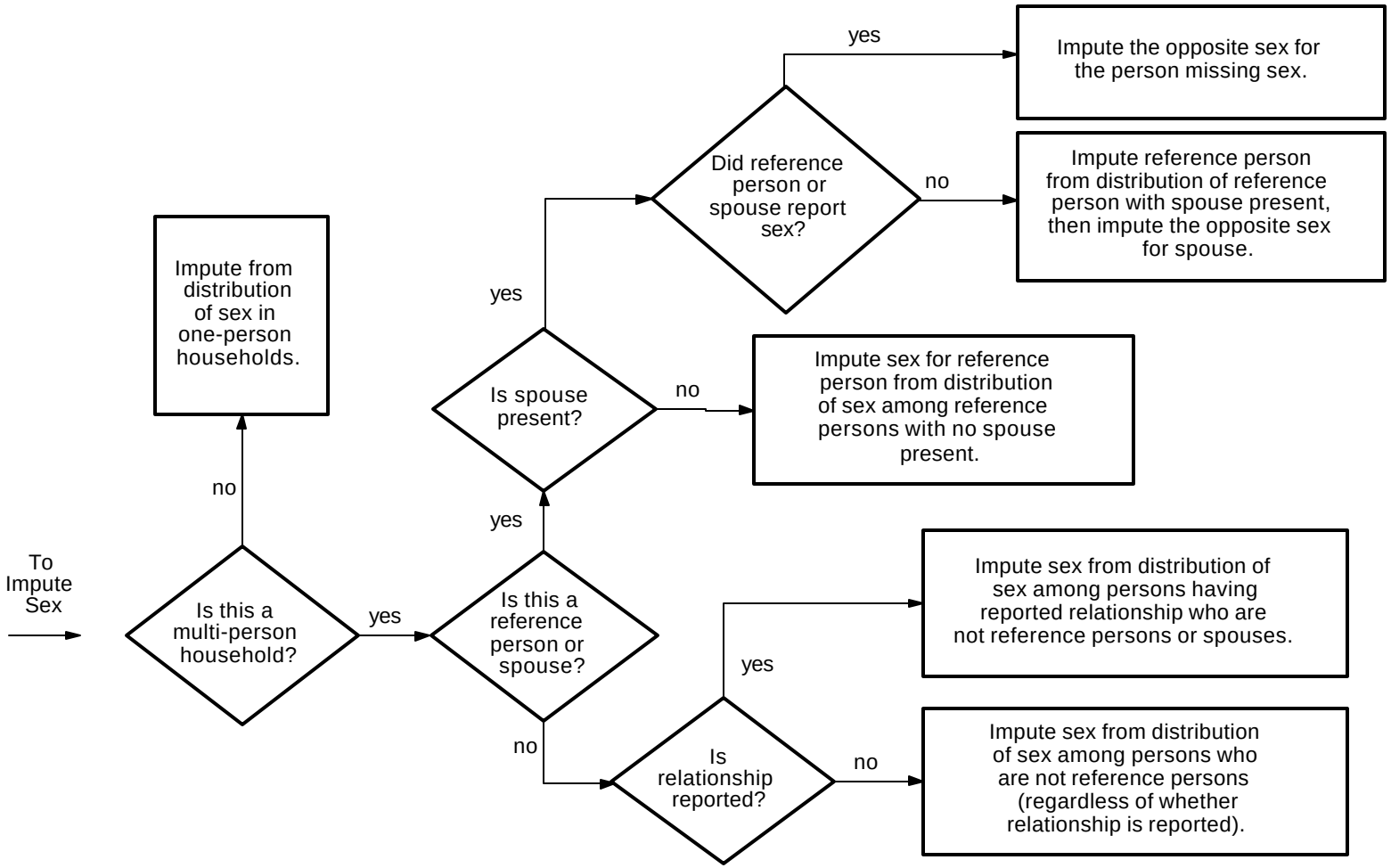
(2) to the person weights of in-movers,
A.C.E. interview-day NI adjustment = 700 / 500 = 1.4



Age is collapsed into the following categories for the purpose of imputation:
0 - 17, 18 - 29, 30 - 49, 50 and over.

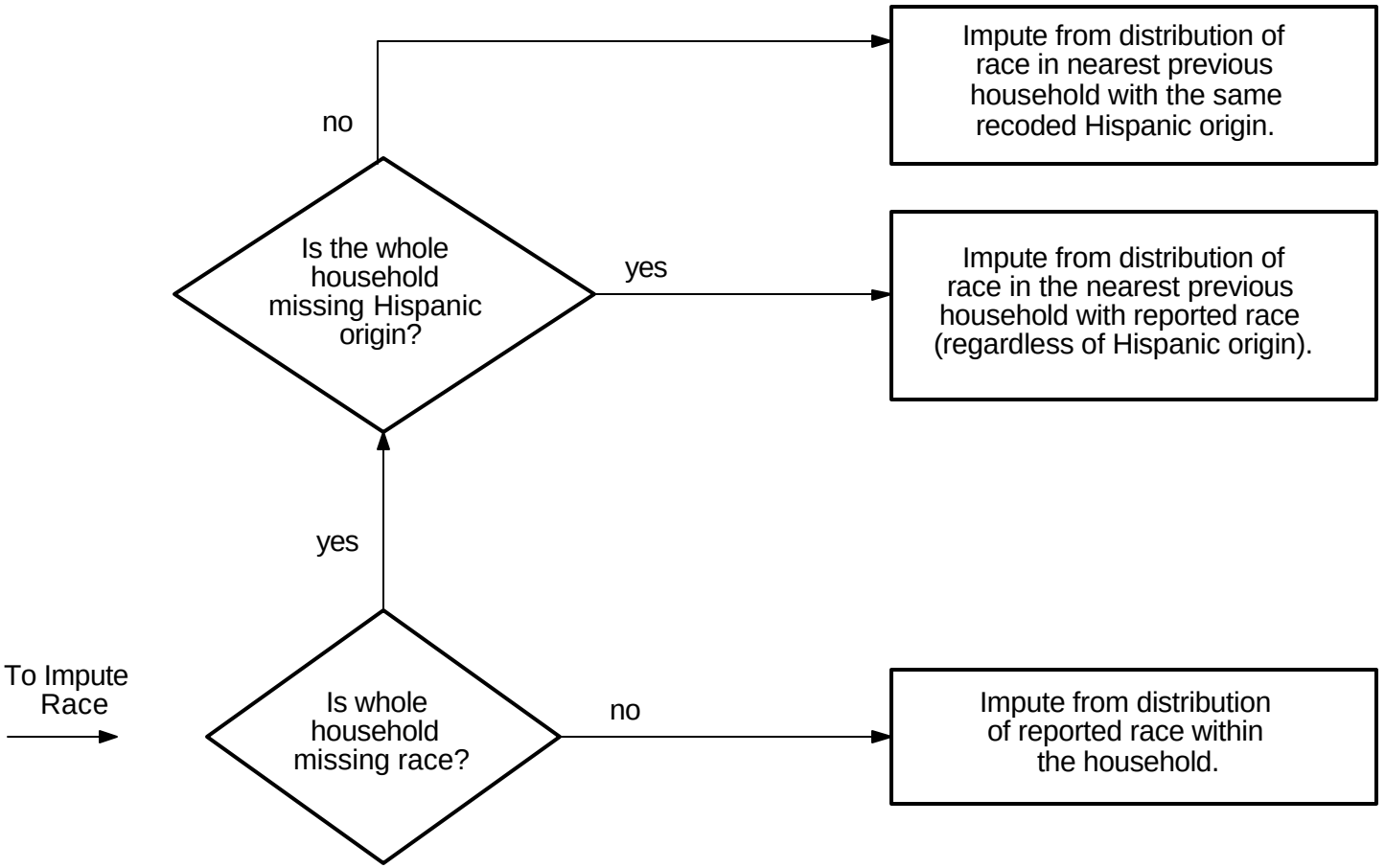
Relationship is collapsed into the following seven categories for the purpose of age imputation:
reference person, spouse, child, sibling, parent, other relative, and nonrelative.

Unless otherwise stated, all distributions are among multi-person households



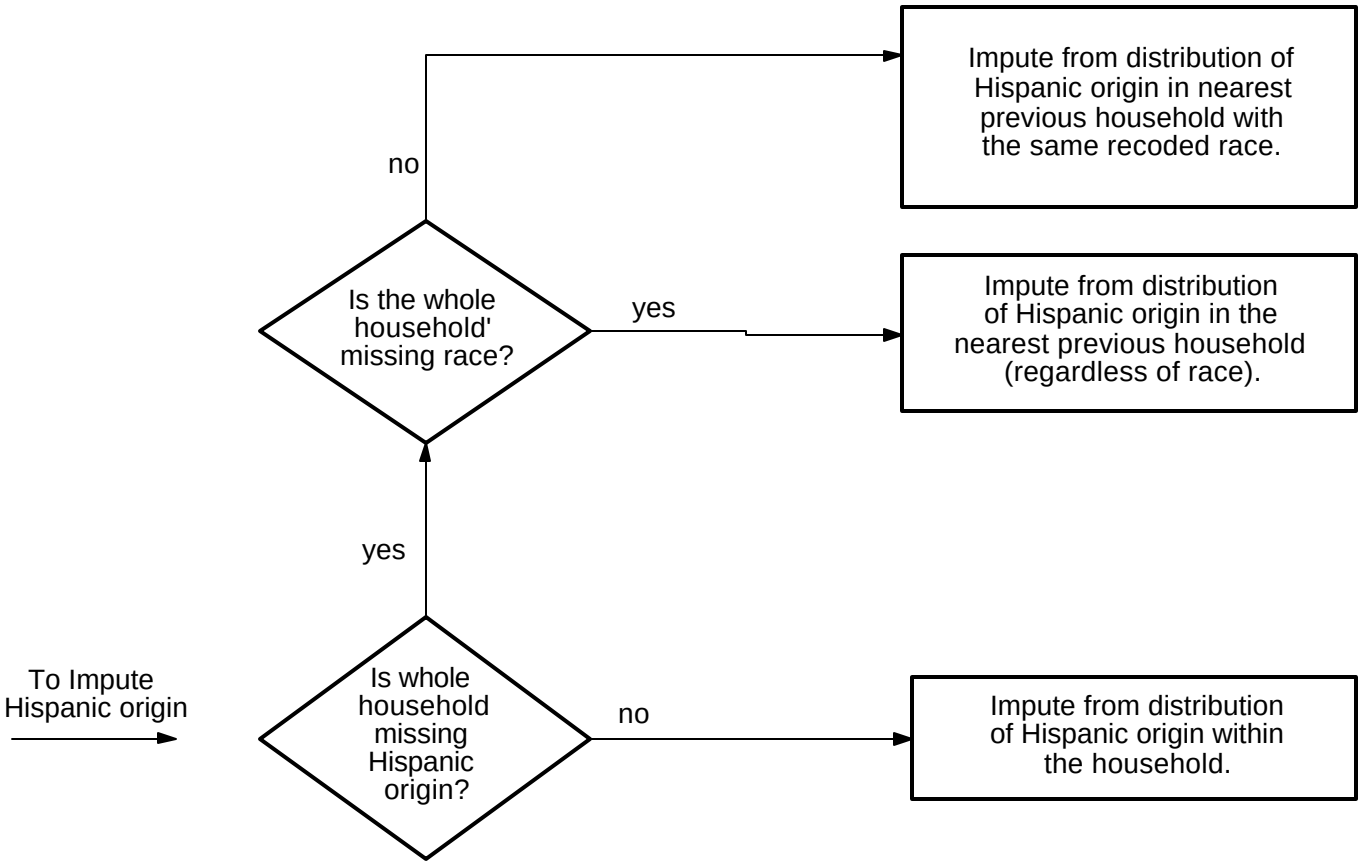
Relationship is collapsed into the following seven categories for the purpose of sex imputation: reference person, spouse, child, sibling, parent, other relative, and nonrelative.

Unless otherwise stated, all distributions are among multi-person households



Note: We can impute any of the 63 possible values of race.

A household's recoded Hispanic origin is the Hispanic origin of the first person on the household roster with reported Hispanic origin. Recoded Hispanic origin has three categories: Hispanic, Non-Hispanic, and missing (indicating the whole household is missing Hispanic origin).



Note: We collapse Hispanic origin to Non-Hispanic/Hispanic for imputation purposes.

A household's recoded race is the race of the first person on the household roster with reported race. Recoded race has four categories: (1) white only, (2) other or both white and other, (3) all reported races not covered by (1) or (2), and (4) missing (indicating that the whole household is missing race). Note that "other" here refers to a race written in by the respondent.

Imputation Cells for Unresolved Cases

Enumeration Status (E Sample)

United States

BFU Group	0 Imputes		1+ Imputes		Total Projected Frequency [1]
1 = Matches needing FU					4,600 (0.6%)
2 = Possible matches					4,600 (0.6%)
3 = Partial HH nonmatches	V3a [2]	V3b	V3a	V3b	35,000 (4.7%)
4 = Whole HH nonmatches (where HU matched); not conflicting HHs	nonhispanic white	other			48,000 (6.4%)
5 = Nonmatches from conflicting HHs; HU not in regular NRFU					
6 = Nonmatches from conflicting HHs; HU in regular NRFU					
7 = Whole HH nonmatches; HU did not match during HU matching	nonhispanic white	other			25,000 (3.4%)
8 = Persons resolved before FU	nonhispanic white	other			611,000 (81.5%)
9 = Persons with insufficient information for matching	(no unresolved cases)				21,600 (2.9%)
Total	600,000 (75 - 85 %)		150,000 (15% - 25 %) [3]		750,000 (100%)

[1] Based on the Dress Rehearsal (DR).

[2] V3a = age 18-29, relationship is child of HHer; V3b = other. In the DR about 12% of BFU Group 3 cases fell into the V3a cell.

[3] Based roughly on prior tests (15% in the Dress Rehearsal w/o Menominee; 20% in 1995; higher in Chicago, 1996).

Puerto Rico

BFU Group	Total Projected Frequency
1 = Matches needing FU	225 (0.6%)
2 = Possible matches	225 (0.6%)
3 = Partial HH nonmatches	1,750 (4.7%)
4 = Whole HH nonmatches (where HU matched); not conflicting HHs	2,400 (6.4%) (in 4, 5, and 6 combined)
5 = Nonmatches from conflicting HHs; HU not in regular NRFU	
6 = Nonmatches from conflicting HHs; HU in regular NRFU	
7 = Whole HH nonmatches; HU did not match during HU matching	1,275 (3.4%)
8 = Persons resolved before FU	30,550 (81.5%)
9 = Persons with insufficient information for matching	1,100 (2.9%) (no unresolved cases)
Total	37,500 (100%)

Imputation Cells for Unresolved Cases

Match Status (P Sample)

Note 1: In the Dress Rehearsal, about 96% of all unresolved matches were people with insufficient information for matching. These cases are *not* followed up. 90% or more of them did not have a valid name; their imputation rates were higher than other cases for race (27.3%), hispanic origin (21.6%), sex (26.2%), age (45.6%), and probably tenure. (Data from Menominee are excluded from these rates.)

Note 2: In both tables, to get a rough idea of the frequencies, it is assumed that all characteristics are independent; we realize that the characteristics may be strongly correlated. Further, the proportions and correlations for Puerto Rico are likely to be different from those of the U.S.

United States

Mover Status	Address Code				Total
	1 = HU Match from initial matching		2 = HU Nonmatch or conflicting household		
Non-mover	0 imputes 601,000	1+ imputes 19,000 [1]	0 imputes 90,000	1+ imputes 3,000 [1]	713,000 (95%)
Mover	0 imputes 32,000	1+ imputes 1,000 [1]	5,000		38,000 (5% [2])
Total	653,000 (87%)		98,000 (13% [3])		750,000

[1] Frequency of 3% - 4% suggested by data from the Dress Rehearsal.

[2] Suggested by Dress Rehearsal data; may be higher due to added time before A.C.E. field operations.

[3] Suggested by current (2000 A.C.E.) HU matching, early results.

Puerto Rico

Mover Status	Address Code		Total
	1 = HU Match from initial matching	2 = HU Nonmatch or conflicting household	
Non-mover	31,000	4,625	35,625 (95%)
Mover	1,625	250	1,875 (5% [2])
Total	32,625 (87%)	4,875 (13% [3])	37,500

Imputation Cells for Unresolved Cases

Residence Status (P Sample)

Note 1: Residence status was treated differently in the 1990 PES, because mover procedure B was used.

Note 2: In the U.S. and Puerto Rico, for people with insufficient information for matching (BFU Group 7): within tenure (× race groups, for the U.S.), derive the weighted residence proportion across BFU Groups 1 - 5; that is, discount cases in Group 6 (those resolved before follow-up)

United States

BFU Group	Owner				Non-Owner				Total Projected Frequency [1]
	Nonhispanic White		Other		Nonhispanic White		Other		
1 = Matches needing FU									4,800 (0.6%)
2 = Possible matches									4,800 (0.6%)
3 = Partial HH nonmatches	V3a	V3b	V3a	V3b	V3a	V3b	V3a	V3b	35,600 (4.7%)
4 = Whole HH nonmatches needing FU (not conflicting HHs)									24,400 (3.3%)
5 = Nonmatches from conflicting HHs needing FU									
6 = Persons resolved before FU									673,000 (89.7%)
7 = Persons with insufficient information for matching (note: no resolved cases)	weighted average over BFU groups 1 - 5		weighted average over BFU groups 1 - 5		weighted average over BFU groups 1 - 5		weighted average over BFU groups 1 - 5		7,600 (1.0%)
Total									750,000 (100%)

[1] Based very roughly on Dress Rehearsal sites, Sacramento and South Carolina

Puerto Rico

BFU Group	Owner	Non-Owner	Total Projected Frequency [1]
1 = Matches needing FU			225 (0.6%)
2 = Possible matches			225 (0.6%)
3 = Partial HH nonmatches			1,750 (4.7%)
4 = Whole HH nonmatches needing FU (not conflicting HHs)			1,250 (3.3%)
5 = Nonmatches from conflicting HHs needing FU			
6 = Persons resolved before FU			33,650 (89.7%)
7 = Persons with insufficient information for matching (note: no resolved cases)	wgt. avg. over BFU groups 1 - 5	wgt. avg. over BFU groups 1 - 5	375 (1.0%)
Total			37,500 (100%)

ESCAP MEETING NO. 10 - 06/28/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 10

June 28, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The tenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on June 28, 2000 at 10:30. The agenda for the meeting was A.C.E. missing data.

Persons in attendance:

Kenneth Prewitt
Nancy Potok
Paula Schneider
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Sally Obenski
John Long
Susan Miskura
Raj Singh
Tommy Wright
Patrick Cantwell
Rita Petroni
Gregg Robinson
Carolee Bush
Maria Urrutia
Annette Quinlan

I. A.C.E. Missing Data

This was the detailed presentation to the ESCAP on missing data. The missing data procedures have been finalized and the purpose of the meeting was to update the ESCAP.
Pat Cantwell presented the results and a summary document is attached. The detailed missing data procedures may be found in the DSSD Memorandum Series Chapter Q-25.

The highpoints of the missing data discussion were as follows:

- Missing data may occur in three areas of the A.C.E.: noninterviews, item nonresponse, and status.

- They are addressed using the following basic methodology:

  (1) Noninterviews are handled through two weighting adjustments, one applied to in-movers and the other to out-movers and non-movers.

  (2) Item nonresponse is addressed through imputation. We use a hot deck approach to impute for tenure, race, and Hispanic origin. Age category and sex are imputed using distributions based on responses.

  (3) Three types of status can remain unresolved even after all person DSE follow-up is complete:

  a) Enumeration status for E-Sample persons - whether the person was correctly or erroneously enumerated in the census.
  b) Match status for P-Sample persons - whether the person matched to someone enumerated in the census.
  c) Residence status for P-Sample people - whether the person was a resident at that address on Census Day.

  For people with unresolved status, we use an imputation cell procedure whereby resolved and unresolved persons are allocated to cells according to their operational or other characteristics. Unresolved persons are given a probability of Enumeration, Match, or Residence status equal to the weighted proportion among the resolved cases in the cell.

  In 1990 we used a logistic regression modeling approach to address cases with unresolved status. This approach also resulted in the assignment of probabilities for the categories described above. We made this change because the cell method will perform adequately for the assignment of probabilities, and offers the advantages of operational efficiency and more straightforward validation.

II. Next Meeting

The next meeting scheduled for Wednesday July 12, 2000 will discuss Correlation Bias.

ESCAP Committee

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Kenneth Prewitt
Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogen
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 11 - 07/12/00

AGENDA

Kathleen P Zveare
07/11/2000 02:39 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC
cc: J Gregory Robinson/POP/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC
Subject: Agenda for 7/12 ESCAP Meeting

The agenda for the July 12 ESCAP Meeting is as follows:

1. Overview of Correlation Bias---Howard Hogan

2. How Do We Measure Correlation Bias Using DA---Greg Robinson

3. Correlation Bias for Evaluation Purposes--Rita Petroni

ESCAP MEETING NO. 11 - 07/12/00

HANDOUTS

Presentation on Demographic Analysis and Measurement of "Correlation Bias"

July 12, 2000

- **The Method of Demographic Analysis**

  Examples of components
  Historical record of DA coverage measurements

- **Differences of DA and Survey coverage results that point to "correlation bias"**

  Differences in net undercount rates
  Differences in sex ratios

- **Challenges for DA coverage measurement in Census 2000**

  Measurement of uncertainty
  Accommodation to new race question (mark one or more)
  Greater focus on sex ratios

What is Demographic Analysis?

, Population (<65)= Births (since 1935)
- Deaths (to persons born after 1935)

+Immigrants (born after 1935)
- Emigrants (born after 1935)

, Population (65+)= Medicare Count
+ Estimated unenrolled

Table 1: Illustrative Values of DA Components for the Estimated U.S. Resident Population, April 1, 2000 (Numbers in Millions)

Component	Age in 2000				
	All Ages	Under 15	15-44	45-64	65+
Total	280.0	60.0	117.0	68.1	35.0
Under age 65:					
Births	235.0	59.0	112.0	64.0	-
Deaths	-14.0	-0.7	-5.0	-8.3	-
Immigrants	29.0	2.0	12.0	15.0	-
Emigrants	-5.0	-0.3	-2.0	-2.7	-
Ages 65+:					
Medicare	35.0	-	-	-	35.0

Figure 1. Percent Net Undercount by Race
1940-1990

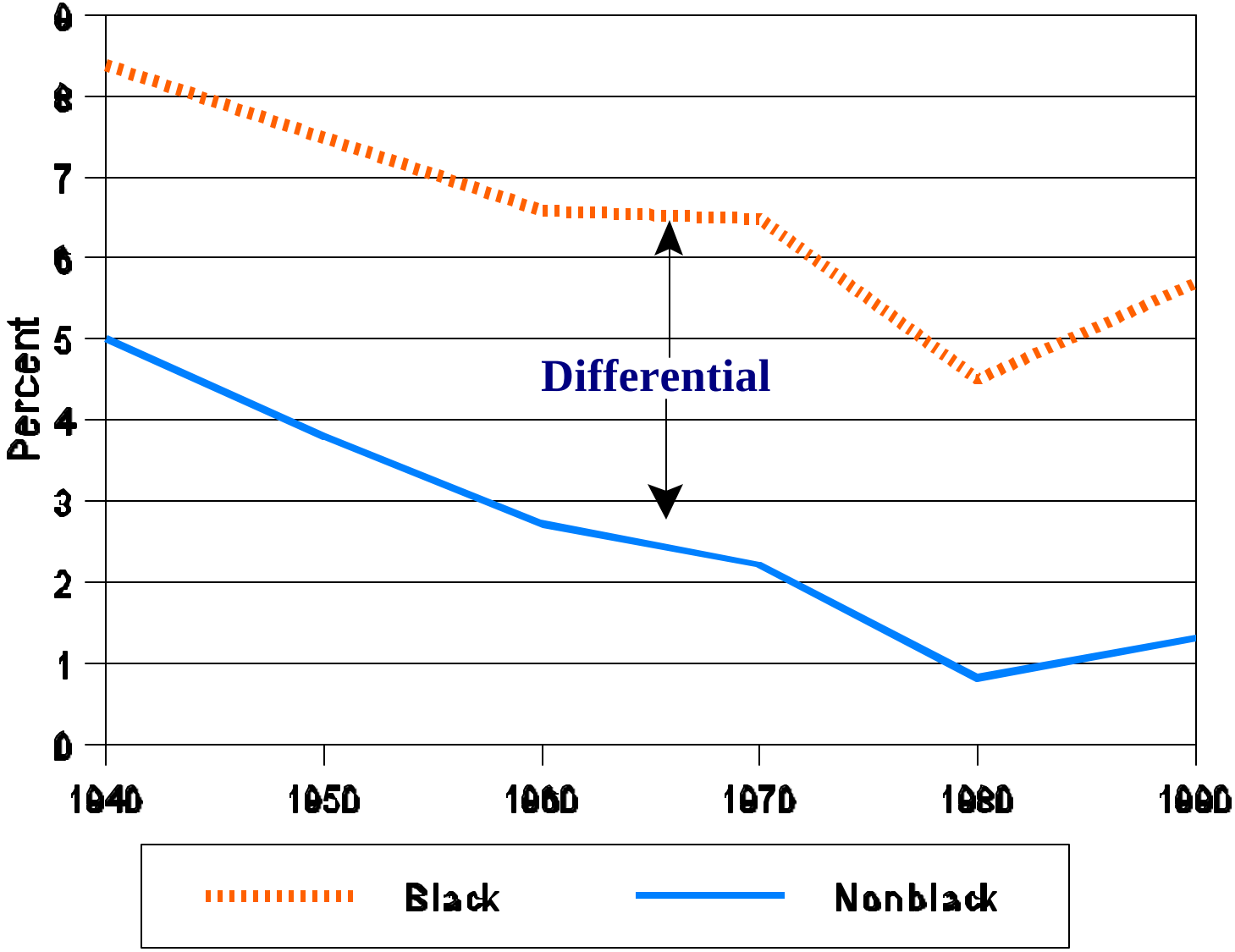
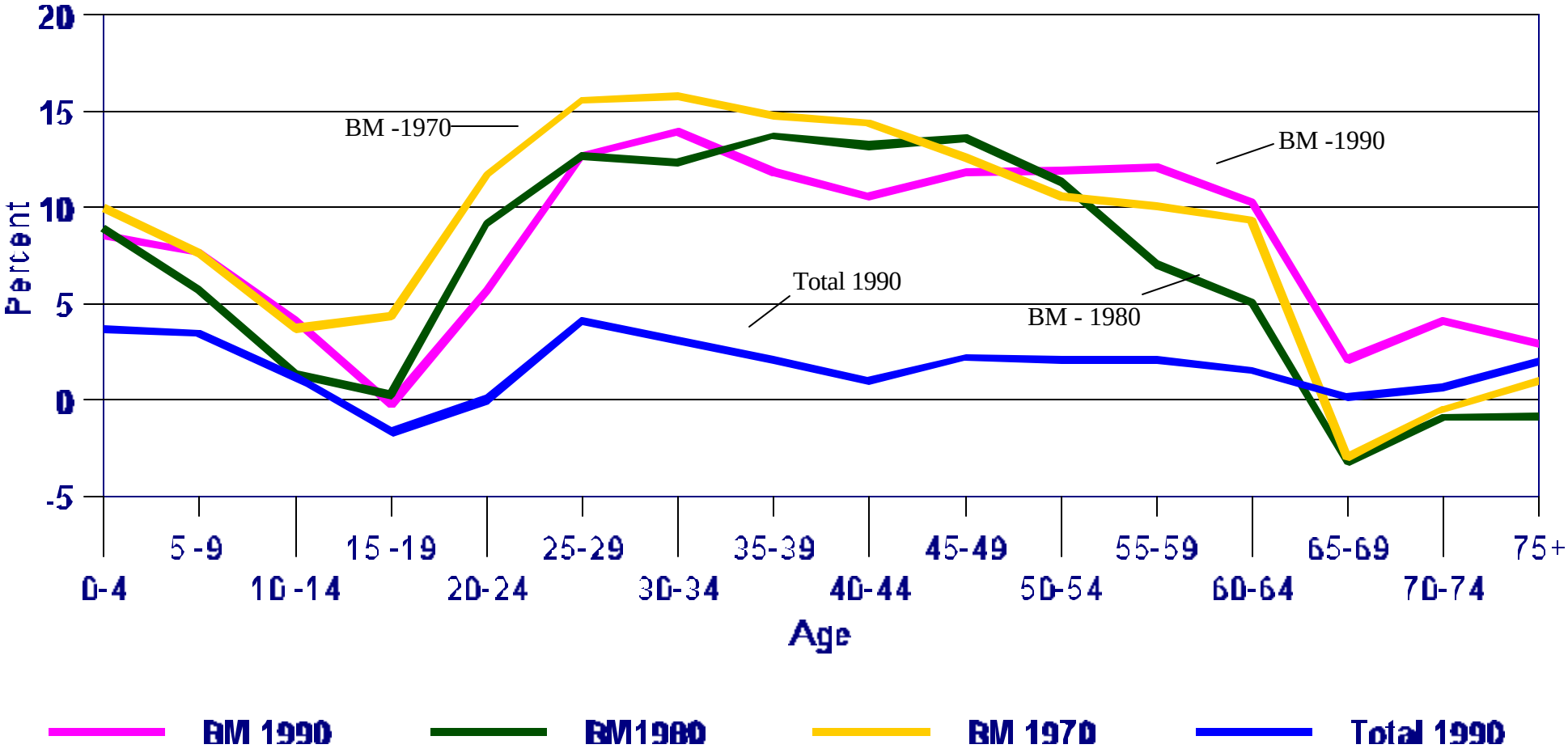
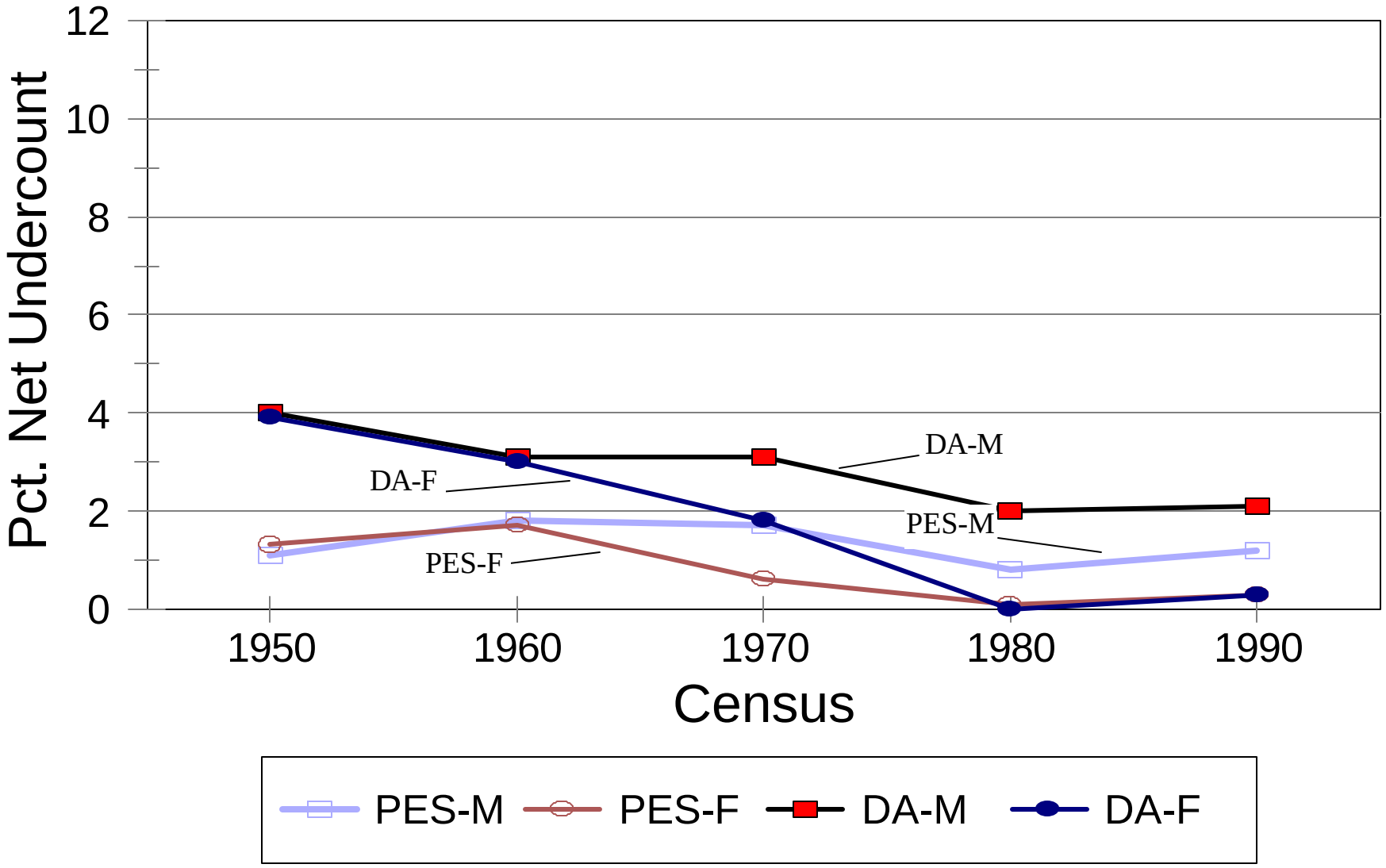


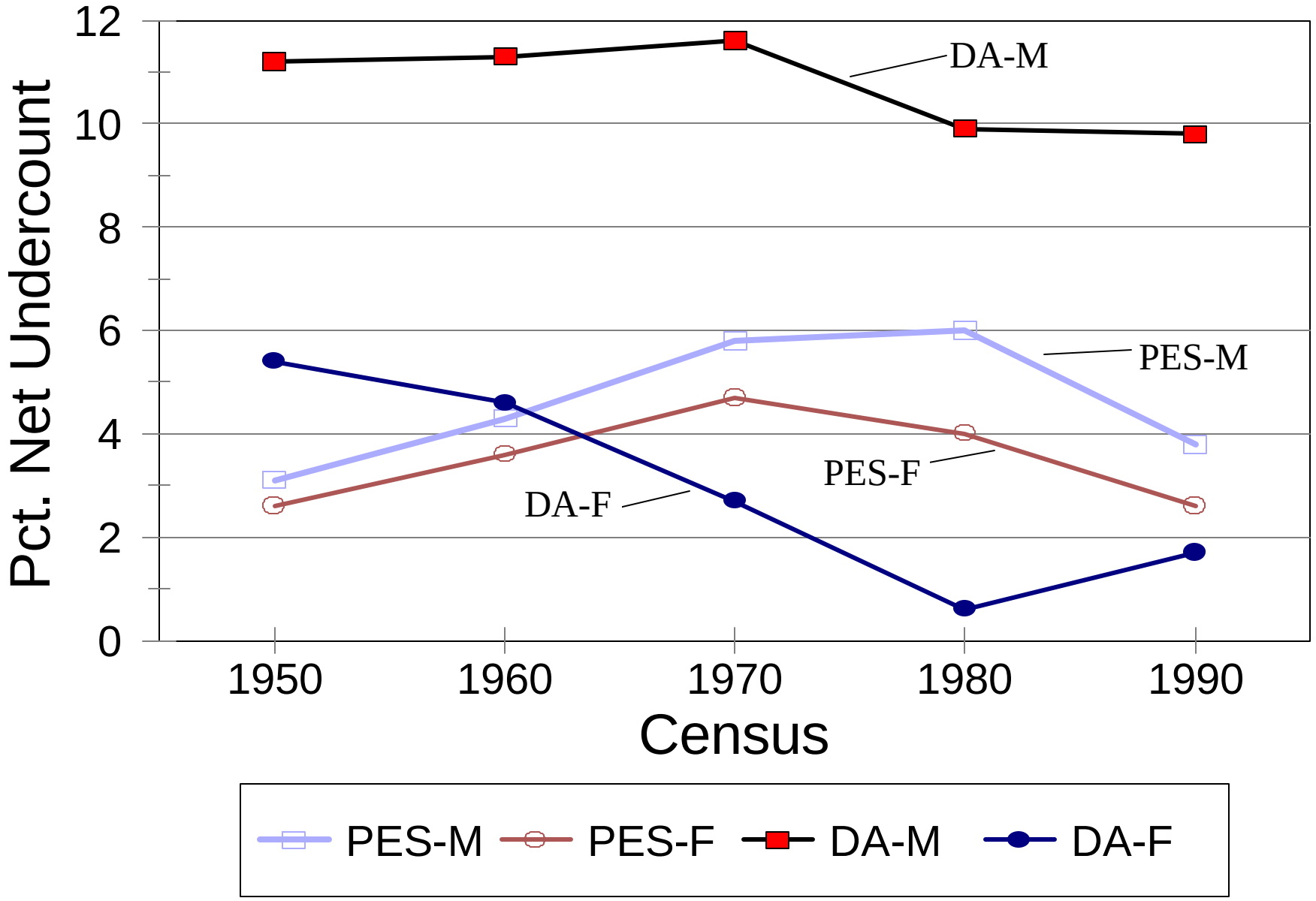
Figure 2. Percent Net Undercount: Black Males 1970 - 1990, and Total, 1990



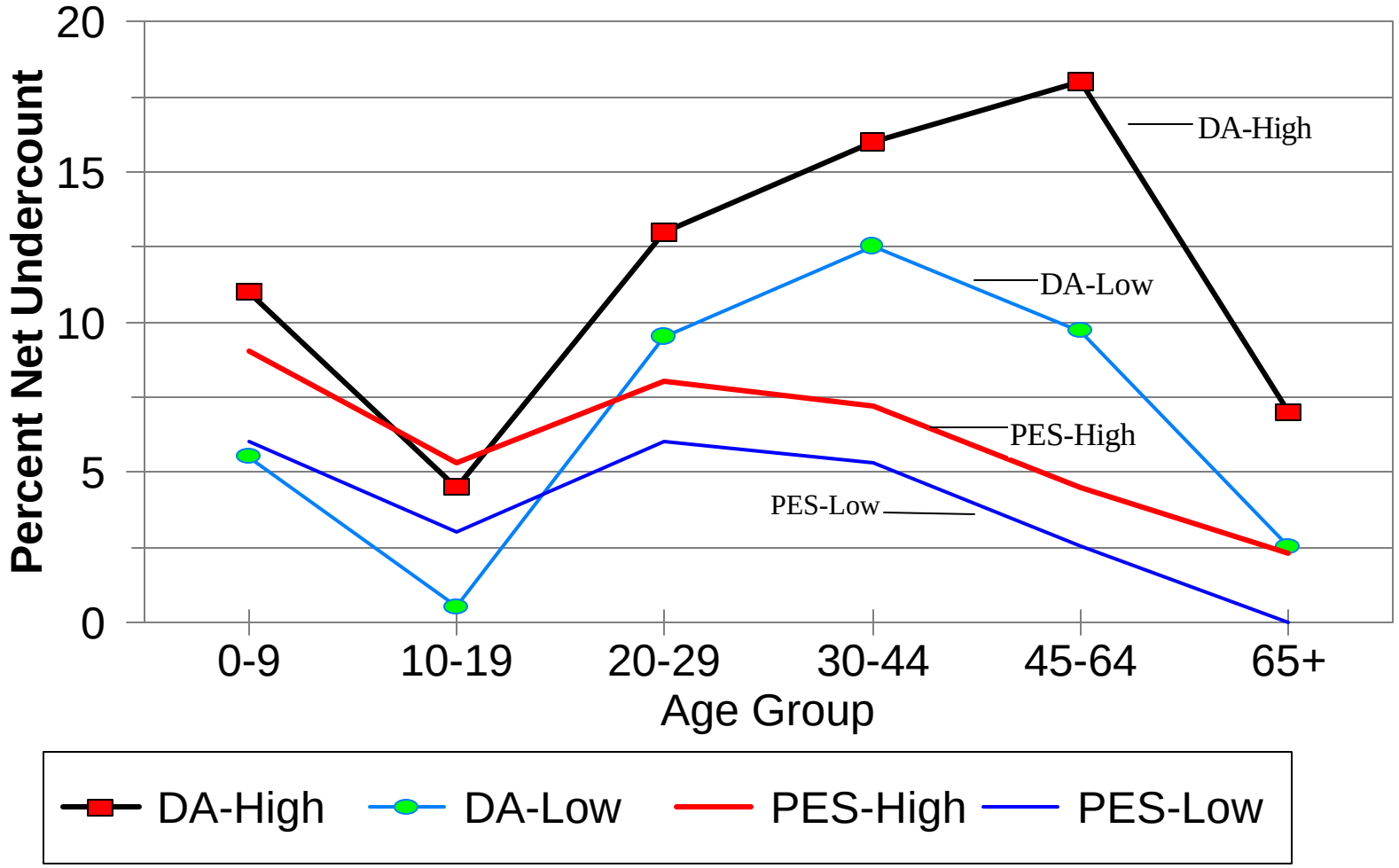
DA and Survey Estimates of Net Undercount for Nonblack Adults, by Sex



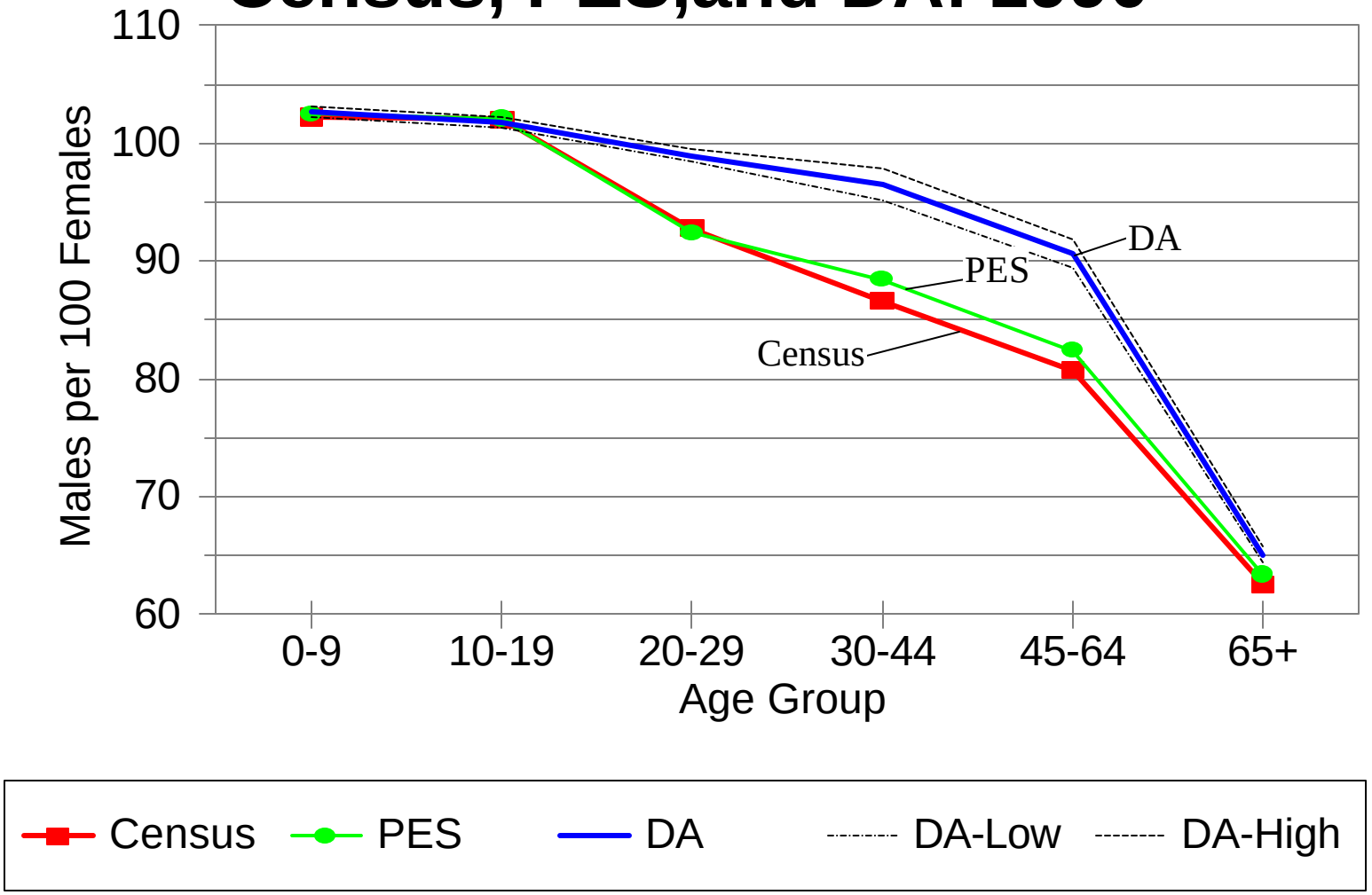
DA and Survey Estimates of Net Undercount for Black Adults, by Sex



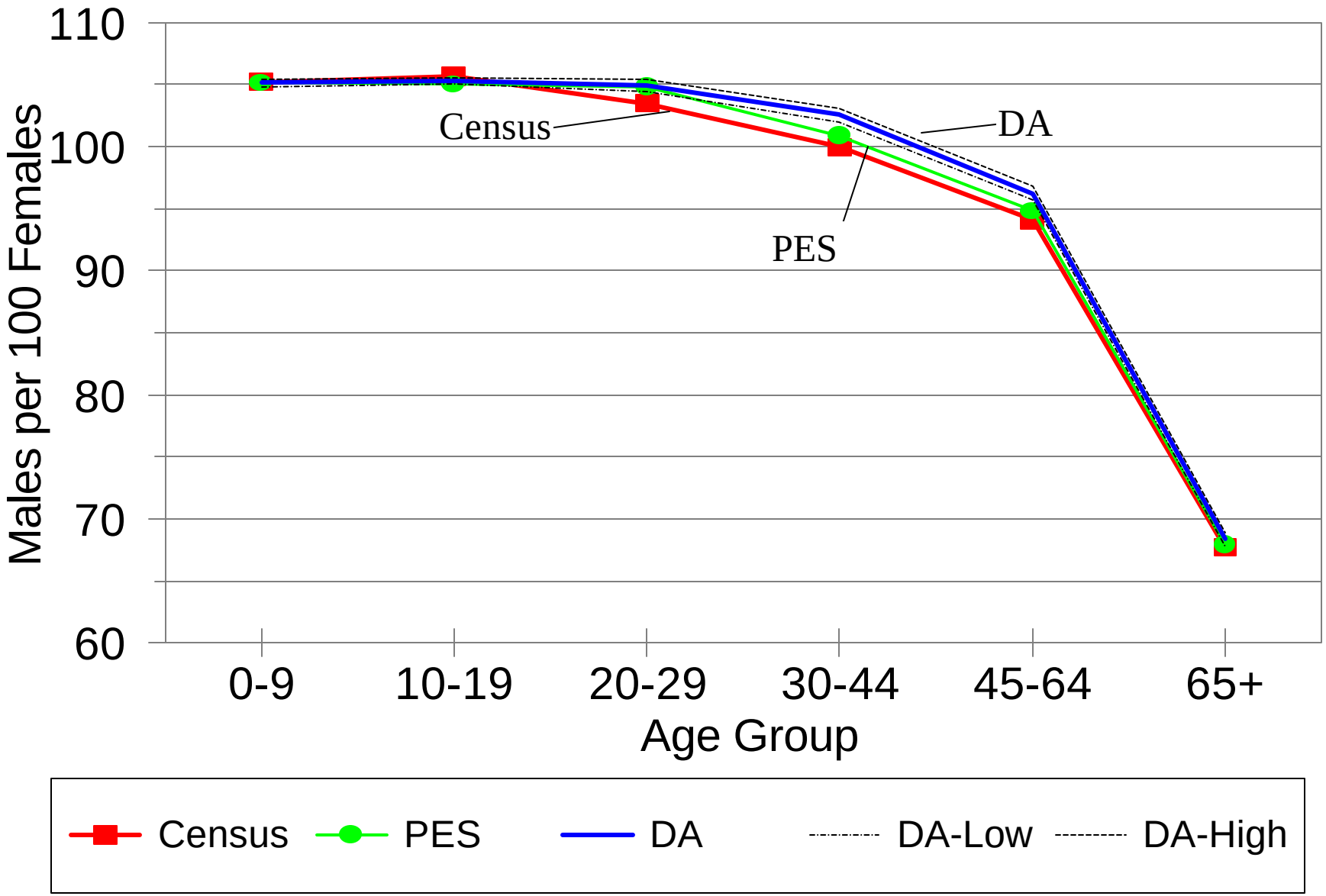
1990 Undercount Uncertainty Intervals
Black Males: DA and PES



Comparison of Sex Ratios for Blacks Census, PES,and DA: 1990



Comparison of Sex Ratios for Nonblacks Census, PES,and DA: 1990



Sex Ratios for Black Adults
Census, Surveys, and DA: 1960-1990

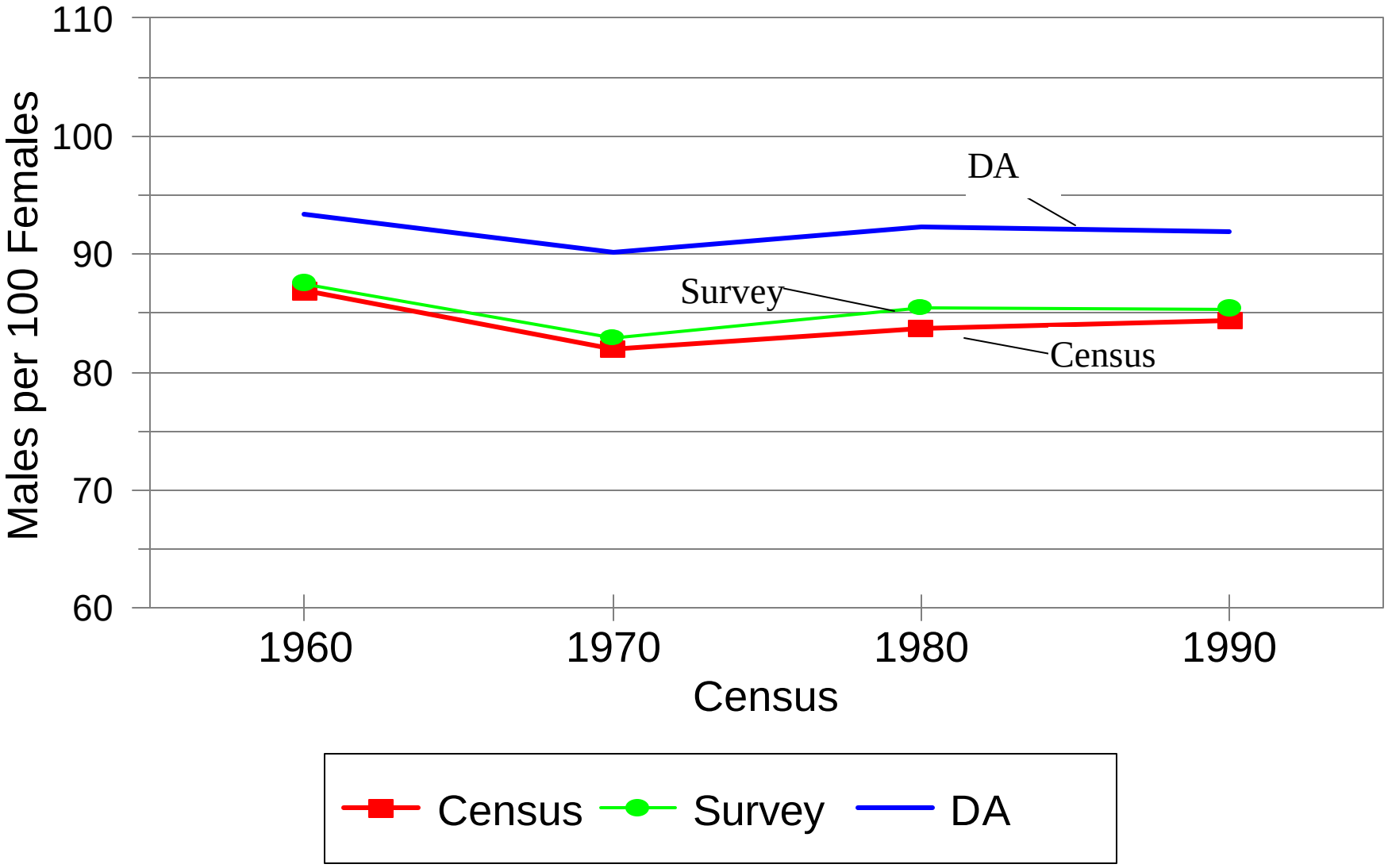
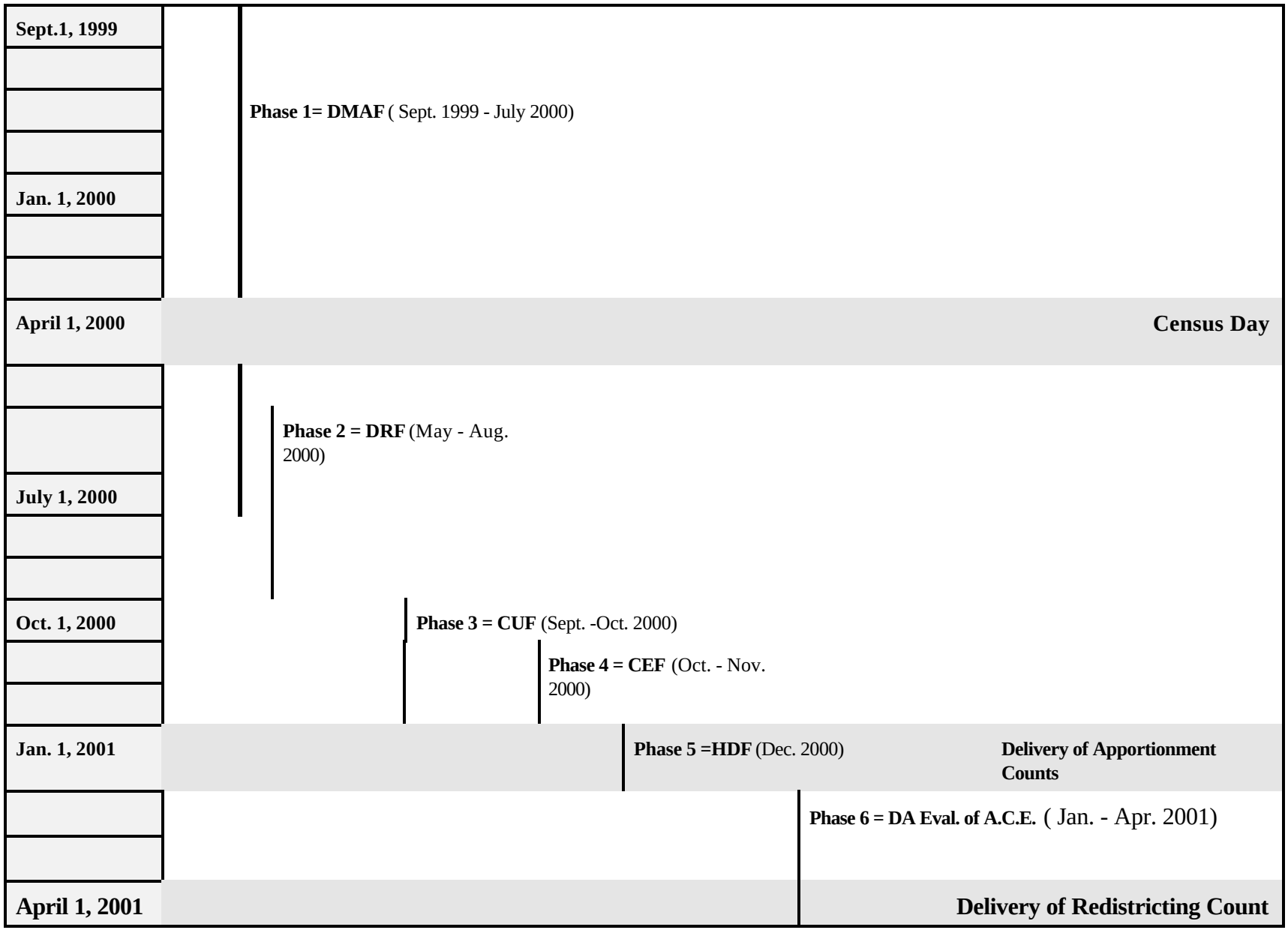


Figure 10
Sequence of Activities for Census 2000 Demographic Analysis Program



November 3, 1997

MEMORANDUM FOR: Distribution List

FROM: J. Gregory Robinson
Chief, Population Analysis and Evaluation Staff

SUBJECT: **The Differential Undercount of Adult Black Men: Is it a Myth?**

A major goal of Census 2000 is to reduce the differential undercount. The most widely recognized differential is the chronic high undercount of adult black men. For the past six census the undercount rate of adult black males has ranged from 10 to 12 percent. These rates are approximately 8 percentage points higher than the overall undercount.

This statement is made from the perspective of Demographic Analysis (DA). Past surveys, including the 1990 PES, have detected coverage differences but not nearly to the extent of DA. If we want a "one number census" that is demographically sound, then steps must be taken to ensure that the survey estimates of undercount approximate the levels suggested by DA.

The Problem

We plan to use Dual System Estimates (DSE) as a vehicle for eliminating differential undercounts in the 2000 census. However, historical survey estimates give little assurance that the large undercounting of Black men will be observed or remedied. As demonstrated below, the survey approach has never measured an undercount of adult black men that is significantly higher (both statistically and realistically) than that of black women or black children.

Table 1. **Percent Net Undercount: Survey Estimates**

		Black	Black	Black Adults	
Year	Total	Total	Children	Male	Female
1990	1.6	4.4	7.0	**3.8**	2.6
1980	1.4	6.0	7.6	**6.0**	4.0
1970	2.3	6.2	7.8	**5.8**	4.7
1960	1.8	3.6	3.1	**4.3**	3.6
1950	1.4	3.2	3.9	**3.1**	2.6

Source: Appendix Table 1

The 1990 PES is a case in point. While the total Black undercount of 4.4 percent was significantly higher than the overall total of 1.6 percent, the Black undercount was most pronounced for children (7.0 percent)--not adult men. The adult Black male (3.8) and female (2.6) percent undercounts were

lower and the male-female difference is not statistically significant.

Where did the media and other observers arrive at the conclusion that black men are severely undercounted? The answer is its simply the profile of net undercount measured by demographic analysis:

Table 2. **Percent Net Undercount: DA Estimates**

		Black	Black	Black Adults	
Year	Total	Total	Children	Male	Female
1990	1.8	5.7	5.9	**9.8**	1.7
1980	1.2	4.5	3.7	**9.9**	0.6
1970	2.7	6.5	5.9	**11.6**	2.7
1960	3.1	6.6	5.0	**11.3**	4.6
1950	4.1	7.5	6.4	**11.2**	5.4
1940	5.4	8.4	7.5	**12.0**	6.5

Source: Appendix Table 1

DA shows a disproportionately high undercount for adult black males--9.8 percent to 12.0 percent nationally (!) in every census since 1940. The undercount rate of children is also relatively high, but only one-half the size of the black male undercount. The net undercounts of black women are moderate in comparison.

There are two reasons why the survey estimates show such a different pattern for adults. The persistent understatement of black men is attributable to the "correlation bias" problem. That is, many persons missed by the census are not being picked up in the survey interview, leading to an understatement of the measured undercount. The overstatement of the net undercount of black women (relative to DA) is less well-known and understood.

Appendix Table 1 provides additional detail on the net undercount estimates for Nonblacks. The DA and survey estimates do not differ as much as for Blacks. In fact, the two approaches measure the same small undercount for adult nonblack females. However, the survey results consistently understate the undercount of adult nonblack men (relative to DA).

A Final Point: Apart from compensating for correlation bias, DA provides a basis to "smooth" age-sex anomalies in the survey estimates. The most obvious example is where the 1990 PES measured a larger undercount of Black women aged 18-29 (5.5 percent) than Black men (3.6 percent). That result is completely contradicted by the 1990 DA estimates (Black female = 2.9, males = 7.7) and the historical record. The detailed age structure of the PES estimates also failed consistency standards.

The Solution: Incorporate Demographic Analysis (DA)

Demographic analysis is the standard for describing historical trends in coverage and differentials by age, sex, and race. The current census 2000 plan--that relies on survey estimation alone for ICM--runs the risk of failing to reduce the differential undercount in a demographically consistent manner. In particular, DA could expose the failure of the 2000 census to reduce the adult Black male undercount. The incorporation of DA into the ICM process can help ensure our goal of a "one number census that is right the first time". If DA is not used, we must spell out the specific improvements to the survey methodology that will render moot the need for DA.

cc:

Schneider (Dir)
Thompson
Bounpane
Marx
Waite
Fay
Mckenney

Hogan (SVSD)

Killion (DSSD)
Vacca
Whitford
Waltman
Griffin
Schindler
Haines

Wright (SRD)
Singh
Bell
Isaki
Petroni
Weiler (FLD)
Blass

Long (POP)
Wetrogan
Miller
Del Pinal
Spencer
Hollmann
Das Gupta
West
Word
Robinson

APPENDIX TABLE 1: COMPARISON OF ESTIMATES OF PERCENT NET UNDERCOUNT BASED ON SURVEY AND DEMOGRAPHIC MEASUREMENT APPROACHES: 1940 - 1990

Census	Coverage Evaluation Program	Total			Black			Nonblack		
		Both Sexes	Black	Non-black	Children	Adults Male	Adults Female	Children	Adults Male	Adults Female
		(1)	(2)	(3)	(4)	(5)	(6)		(7)	(8)
A. Survey approach										
1990	Post Enumeration Survey	1.6	4.4	1.2	7.0	3.8	2.6	2.5	1.2	0.3
1980	Post Enumeration Program	1.4	6.0	0.8	7.6	6.0	4.0	1.5	0.8	0.1
1970	CPS/Census Match	2.3	6.3	1.8	7.8	5.8	4.7	2.1	1.7	0.6
1960	Survey Coverage Study	1.8	3.6	1.6	3.1	4.3	3.6	1.2	1.8	1.7
1950	Post Enumeration Survey	1.4	3.2	1.2	3.9	3.1	2.6	1.1	1.1	1.3
1940	N/A									
B. Demographic approach										
1990	Demographic Analysis	1.8	5.7	1.3	5.9	9.8	1.7	1.6	2.1	0.3
1980	Demographic Analysis	1.2	4.5	0.8	3.7	9.9	0.6	0.3	2.0	0.0
1970	Demographic Analysis	2.7	6.5	2.2	5.9	11.6	2.7	1.8	3.1	1.8
1960	Demographic Analysis	3.1	6.6	2.7	5.0	11.3	4.6	2.0	3.1	3.0
1950	Demographic Analysis	4.1	7.5	3.8	6.4	11.2	5.4	3.3	4.0	3.9
1940	Demographic Analysis	5.4	8.4	5.0	7.5	12.0	6.3	6.3	4.9	4.7
C. Method difference (= A - B)										
1990	Survey minus DA	-0.2	-1.3	-0.1	1.1	-6.0	0.9	0.9	-0.9	0.0
1980	Survey minus DA	0.2	1.5	0.0	3.9	-3.9	3.4	1.2	-1.2	0.1
1970	Survey minus DA	-0.4	-0.2	-0.4	1.9	-5.8	2.0	0.3	-1.4	-0.2
1960	Survey minus DA	-1.3	-3.0	-1.1	-1.9	-7.0	-1.0	-0.8	-1.3	-0.3
1950	Survey minus DA	-2.7	-4.3	-2.6	-2.5	-8.1	-2.8	-2.2	-2.9	-2.6

Note:

Demographic analysis estimates represent percent net undercoverage for all years. Estimates for 1940-80 represent revised estimates that are consistent with the methodology and components used to produce the 1990 demographic estimates.

Survey-based estimates represent percent net undercoverage for all years. Survey estimates for 1960 and 1950 refer to Black-and-other-races instead of Black. Estimates for 1980 represent a composite of 9 sets of estimates (sets 14-20, 14-9, and 14-8 are excluded). Approximate sample sizes of estimates: 1990 - 144,000 interviewed households; 1980 - 84,000 CPS households for P-sample and 110,000 census households for E-sample; 1970 - 45,000 interviewed households; 1960 - about 35,000 households; 1950 - about 25,000 households.

Adult - DA estimates refer to population 18 and over in 1990 and 20 and over in 1950-1980; survey estimates for 1990 (PES) refer to population 18 and over in 1990, 20 and over in 1980 (PEP), and 15 and over in 1950-1970.

ESCAP MEETING NO. 11 - 07/12/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 11

July 12, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The eleventh meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on July 12, 2000 at 10:30. The agenda for the meeting was A.C.E. correlation bias.

Persons in attendance:

Kenneth Prewitt
William Barron
Nancy Potok
Paula Schneider
Nancy Gordon
John Thompson
Jay Waite
Howard Hogan
John Long
Susan Miskura
Donna Kostanich
Fay Nash
Tommy Wright
David Hubble
Rita Petroni
Gregg Robinson
Carolee Bush
Maria Urrutia
Annette Quinlan

I. A.C.E. Correlation Bias

The purpose of this meeting was to describe and discuss correlation bias with the ESCAP. Howard Hogan provided a general discussion of correlation bias and Gregg Robinson discussed methods to measure correlation bias based on Demographic Analysis.

Howard Hogan defined the two causes of correlation bias, causal dependence and heterogeneity, the handout is attached. Causal dependence occurs when the event of being included in the census affects a person's chance of being included in the A.C.E., or vice versa. Correlation bias also results from heterogeneity of the probabilities of being included in either the census or the coverage measurement surveys. The theoretical assumptions underlying the DSE do not require that the initial census and the coverage measurement survey have the same probability of including people. However, DSE does assume that there are not groups of the population within post-strata that have different inclusion probabilities for both the initial census and the coverage measurement survey. When this situation occurs the DSE will be subject to correlation bias and will understate the "true" population totals. The attachment describes examples of this phenomenon.

Gregg Robinson provided an overview of the Demographic Analysis (DA) program, which included a description of the DA method, the major finding of DA measurements of net undercount trends and differentials from previous censuses, and how the coverage patterns based on DA estimates compare to coverage patterns measured by previous census coverage measurement surveys (such as the 1990 PES) and how DA provides a measurement of correlation bias. Differences in the DA and DSEs in 1990 and 1980 for some age-sex groups (such as adult Black men between the ages of 18 and 29) provide measures of "correlation bias."

The method of DA relies on aggregate administrative records, which are essentially independent of the census. The DA estimates for the population under 65 years of age in 2000 (born after 1935) are based on the compilation of historical data or estimates of births, deaths, immigrants, and emigrants. Administrative Medicare data are used to estimate the population 65 and over. Limitations of the DA estimates were also discussed, including problems with estimating some of the components (e.g., undocumented immigrants), the inability to provide coverage estimates for detailed race/ethnic groups, and how DA only provides estimates at the national level.

DA has been used over the years to describe historical trends in coverage differentials by age, sex, and race (Black, Nonblack). It has provided a consistent tracking system by which the percent undercount rates can be compared from decade to decade. DA has measured a persistent and disproportionate undercount of adult Black men and Black children in the censuses of 1940 to 1990. The net undercount of adult Black men during the 1940 - 1970 censuses exceeded 10 percent nationally and during the 1980 and 1990 censuses was approximately 10 percent nationally.

A table and figures were also presented which compared the DA and coverage measurement survey results of each census since 1950. While the two methods have been in close agreement regarding the overall net undercount in the most recent censuses (e.g, the PES measured a net undercount of 1.6 percent in 1990; DA measured a slightly higher rate of 1.8

percent), certain differences emerge in the comparison of estimates dis-aggregated by age, sex, and race. In particular, the survey net undercount estimates for adult Black men are substantially lower than the corresponding DA estimates. The difference remains even after accounting for the "uncertainty" in the measured undercounts.

The Demographic Analysis estimates demonstrate that the coverage measurement survey results have consistently understated the undercount of adult Black men. This is the empirical evidence of "correlation bias". DA sex ratios have been shown to be less subject to uncertainty than the DA "point" estimates themselves. Comparison of DA and coverage measurement survey-based sex ratios (ratio of males to females) for adult Blacks further confirm correlation bias.

II. Next Meeting

The next meeting scheduled for Wednesday July 26, 2000 will discuss plans for how correlation bias will be estimated for A.C.E. evaluation purposes.

ESCAP Committee

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Kenneth Prewitt
Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogen
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 12 - 07/26/00

AGENDA

Kathleen P Zveare
07/25/2000 10:53 AM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC
cc: Rita J Petroni/PRED/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC
Subject: Agenda for 7/12 ESCAP Meeting

The agenda for the July 26 ESCAP Meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

1. Correlation Bias for Evaluation Purposes--Rita Petroni

2. A.C.E. Analysis--Howard Hogan/Debbie Fenstermaker

ESCAP MEETING NO. 12 - 07/26/00

HANDOUTS

DRAFT 7/26/00

Summary of A.C.E. Quality Indicators

For discussion purposes: Please note that this plan is in the process of being developed. Changes and refinements should be assumed.

1. Estimation Quality Indicators

   - Census 2000 counts, A.C.E. estimates with SEs, and A.C.E. undercount rates with SEs for total pop and seven race/origin domains
   - DA undercounts and sex ratios from 1940 to 2000
   - Poststratum-level details about the components of the DSE: weighted and unweighted matches, correct enumerations, P- and E-sample population estimates, movers, census counts, census IIs, sample sizes, DSEs, undercounts, CVs
   - Census counts, undercounts, and CVs for aggregated areas/groups, such as states
   - Comparisons with 1990 PES where appropriate

2. Missing Data Quality Indicators

   - Comparison of weighted P-sample, weighted E-sample, and census item missing data rates for poststratification factors: race, Hispanic origin, age, sex, tenure
   - Weighted estimate of noninterviewed P-sample housing units, and weight distribution of interviewed P-sample housing units before and after noninterview adjustment
   - Distribution of P-sample residency status before and after residence status imputation
   - Distribution of P-sample match status and match rates before and after match status imputation
   - Distribution of E-sample enumeration status and CE rates before and after enumeration status imputation

3. A.C.E. Interviewing Quality Indicators

   - Overall interview results: number of interviews, noninterviews, and non-occupied housing units for interview day and Census Day
   - Detailed interview results: weighted and unweighted number and percentages of household-member interviews, proxies, refusals, other noninterviews, vacants, and non-housing units
   - Distribution of interviews by mode (telephone or personal visit) and mover status
   - NRCO results by census return rate categories, TEA, and MSA status
   - Detailed information about timing of phone and personal visit interview phases by LCO and cumulative completion rate over time by ACERO

4. Person Match Quality Indicators

- After follow-up person match results: matches, A.C.E. non-matches, A.C.E. out-of-scope, census CEs, census EEs, unresolved
- Distribution of match results by race/origin domains

5. P-Sample and E-Sample Weight Quality Indicators

- Trimmed clusters with their net errors before and after trimming
- Distribution of P-sample and E-sample cluster and housing unit weight variation
- Influential clusters identified using a jackknife procedure

6. Sample Design Quality Indicators

- Comparison of weighted P-sample, weighted E-sample, and census housing units and persons
- Comparison of weighted P-sample, weighted E-sample, and census distribution of post-stratum factors
- Comparison of occupancy/vacancy rates for weighed P sample, weighted E sample, and census

7. Other

- Degree of consistency of P-sample and E-sample responses for post-stratification data items, with and without imputed data

Estimation of Correlation Bias for Evaluations

1. What is Correlation Bias?

Dual System Estimation is subject to correlation bias. Correlation bias occurs because people missed in the census enumeration are also more likely to be missed in the A.C.E.

2. What was Done to Measure Correlation Bias in 1990?

For the 1990 Coverage Measurement Evaluations, correlation bias was measured by comparing Post Enumeration Survey (PES) estimates of population size with estimates derived from combining results with demographic analysis (DA) sex ratios according to a method developed by Bell (1993). Sex ratios rather than population totals were used to minimize the effect of errors in demographic analysis.

Bell's approach to estimating correlation bias at the national level:

1. uses the usual dual system estimates (DSE) for females (i.e. the approach assumes no correlation bias).

2. selects a model for males that produces alternative postratum DSEs allowing for some dependence between the census and the coverage measurement survey.

3. estimates the dependence by controlling the alternative male DSEs to reproduce DA sex ratios, assuming female DSEs are correct, when aggregated to the national level.

The national level of correlation bias was then taken to be the difference between the resulting DSE for males and the usual DSE for males. No adjustments were made for "other" sources of bias (Mulry, 1991). The correlation bias was then distributed to the PES poststrata proportional to the estimate of the number of males in the fourth cell of the DSE for the poststratum (CAPE, 1992).

This approach provided the best subnational indications of correlation bias that were available, but it does have several limitations. First, it assumes that demographic analysis sex ratios are accurate and that there is no correlation bias for females. Rough evidence from demographic analysis totals for females in 1980 and 1990 do not refute this latter assumption, but this does not prove that correlation bias for females is entirely absent within poststrata. Also, the suitability of any alternative male DSE for this procedure depends on how will its underlying assumption conforms with reality. Unfortunately, this is uncheckable from our data. A fourth limitation is the occurrence of negative estimates of census counts less the estimates of erroneous enumerations (Bell, 1991). When the estimate is negative, no amount of the estimated people missed due to correlation bias is allocated to that post-stratum (CAPE, 1992). A final potential limitation was not adjusting DSEs for other biases (i.e. measurement bias, contamination bias, ratio estimator bias). This potentially leads to an underestimate of correlation bias.

Spencer (2000c) looked at how unmeasured correlation bias impacts loss function analysis that compares adjusted and unadjusted census counts. He found that if the unmeasured correlation bias is positively correlated with the undercount, the fact that there is unmeasured correlation bias will tend to make the census look unduly favorable relative to the DSE. If unmeasured correlation bias is uncorrelated with the undercount, then there is no systematic favoring toward either the census or the DSE in comparisons of relative accuracy. 1990 PES results and participant observer studies suggest that it is plausible that correlation bias is correlated with the undercount.

3. Models and Assumptions

The 2x2 table used for DSE is:

Census	PES		
	In	Out	Total
In	x_{11}	x_{12}	x_{1+}
Out	x_{21}	$\hat{x}_{22}$	$\hat{x}_{2\%}$
Total	x_{+1}	$\hat{x}_{\%2}$	$\hat{x}_{..}$

For the usual DSE, we assume no correlation bias or that:

$$\frac{\Pr[\text{In PES} \mid \text{In Census}]}{\Pr[\text{In PES} \mid \text{Not In Census}]} = 1.$$

This is equivalent to the assumption that:

$$\frac{x_{11}\hat{x}_{22}}{x_{12}x_{21}} = 1.$$

So Bell's approach assumed for females that:

$$\frac{F_{11}\hat{F}_{22}}{F_{12}F_{21}} = 1.$$

For males, he considered four DSE models based on these assumptions:

1. $$\frac{M_{11}\hat{M}_{22}}{M_{12}M_{21}} = \frac{\Pr[\text{In PES}\mid \text{In Census}]}{\Pr[\text{In PES}\mid \text{Not In Census}]} = C$$

2. $$\frac{M_{11}(M_{21}\%\hat{M}_{22})}{M_{1\%}M_{21}} \approx C$$

3. $$\frac{\hat{M}_{22}\hat{F}_{..}}{\hat{M}_{..}\hat{F}_{22}} \approx C$$

4. $$\frac{M_{21}\hat{M}_{..}}{M_{1\%}(M_{21}\%\hat{M}_{22})} = \frac{\Pr[\text{In PES}\mid \text{Not In Census}]}{\Pr[\text{In Census}]} = C$$

For each model, C was assumed constant within age/sex groups for males.

4. What Does the Census Bureau Plan to Do for Census 2000?

For the 2000 Total Error Model, we again expect to use demographic analysis to obtain national levels of correlation bias and a modeling approach to distribute the correlation bias across poststrata.

The Bureau will re-evaluate whether to use DA data and knowledge of other DSE biases to obtain national levels of correlation bias. We are considering a method based on DA totals and a method based on sex ratios which treats correlation bias for females as negligible. For both methods we are also considering whether to make adjustments for contamination, ratio estimator, and measurement error biases (Spencer, 2000a).

To assist in determination of which estimates of correlation bias to use, we plan to develop point estimates of other biases in DSE estimates of males and females by black and non-black and given resources we may estimate some selected variances.

Measurement error biases will be obtained from an evaluation sample of about 2300 clusters - over

twice the number of evaluation clusters used for the 1990 evaluations.

We are considering various modeling approaches to distribute correlation bias to poststrata. In addition to Bell's approach (1993), we are considering approaches developed by Elliott and Little (1999) and by Haberman, Jiang, and Spencer (1998).

Elliot and Little improve the application of Bell's models. They propose some general principals for aiding the choice among the alternative models. Using these principles, they choose a model and imbed it within a more comprehensive Bayesian model for counts in poststrata of the population. Through judicious choice of parameterization and prior distributions, their Bayesian model eliminates negative cell estimates and reduces outlying predictions of undercount rates. In addition, their model detects, through posterior predictive distributions, strata in which large negative raw cell estimates may be due to bias rather than variance. Their model also allows direct control over the inter-strata sex ratio variation through the variance parameter. Additionally, their method can be extended to provide estimates of precision that incorporate uncertainty in the estimates from demographic analysis and other sources.

The Haberman et.al. approach is similar to the approach of Bell. It is applicable when the capture probabilities can vary from individual to individual according to a logit model, whereas Bell's models apply to the post-stratification model, in which capture probabilities are assumed constant. The two approaches both use sex ratios and in effect constrain the adjustment factors to match the sex ratios. The females can be assumed to have zero correlation bias, or the DA totals can be used to constrain the adjustment factors for both females and males to agree with the DA totals.

The modeling approach we use will be determined by consideration of the strength of each approach and resources available to implement the approaches.

5. References

Bell, W. R. (1991), "Use of Alternative Dual System Estimation to Measure Correlation Bias," 1990 Coverage Studies and Evaluation Memorandum Series #0-3, Project P13, July 10, 1991.

Bell, W.R. (1993), "Using Information from Demographic Analysis in Post-Enumeration Survey Estimates," Journal of the American Statistical Association, 88, 1106-1118.

CAPE (1992), "Assessment of Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates," Report of the Committee on Adjustment of Postcensal Estimates, Bureau of the Census, Department of Commerce, August 7, 1992.

Elliott, M. R. and Little, R. J. A. (1999), "A Bayesian Approach to Combining Information from a

Census, A Coverage Measurement Survey and Demographic Analysis," to appear in the Journal of the American Statistical Association.

Haberman, S. J., Jiang, W. and Spencer, B. D. (1998), "Activity 7: Develop Methodology for Evaluating Model-Based Estimates of the Population Size for States," Final Report under contract 50-YABC-2-66023 for the Bureau of the Census. Chicago: National Opinion Research Center.

Mulry, M.H. (1991), "1990 Post Enumeration Survey Evaluation Project P16: Total Error in PES Estimates for Evaluation Post Strata," 1990 Coverage Studies and Evaluation Memorandum Series #R-6, July 11, 1991. Washington, DC.: Bureau of the Census.

Spencer, B. D. (2000a), "Final Report on Correlation Bias Methodology," Under contract number 50-YABC-7-66020 for the Bureau of the Census. Abt Associates, Inc. and Spencer Statistics, Inc., April 20, 2000.

Spencer, B. D. (2000b), " Total Error Model for Census 2000: How Components of Error Can Be Estimated from the Bureau's Planned Evaluation Studies", Final Report under contract number 50-YABC-7-66020 for the Bureau of the Census. Abt Associates, Inc. and Spencer Statistics, Inc., revised May 17, 2000.

Spencer, B. D. (2000c), " A Strategy for Analyzing Whether Adjustment of Census 2000 Will Improve Redistricting," draft 1.0, June 8, 2000.

ESCAP MEETING NO. 12 - 07/26/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 12

July 26, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The twelfth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on July 26, 2000 at 10:30. The agenda for the meeting was A.C.E. correlation bias for evaluation purposes and A.C.E. quality analysis.

Committee Attendees:

William Barron
Nancy Potok
Cynthia Clark
Nancy Gordon
John Thompson
Howard Hogan
John Long
Susan Miskura

Other Attendees:

Kenneth Prewitt
Donna Kostanich
Fay Nash
Raj Singh
Tommy Wright
David Hubble
Rita Petroni
Gregg Robinson
Maria Urrutia
Annette Quinlan

I. A.C.E. Estimation of Correlation Bias for Evaluations

Rita Petroni presented and distributed the correlation bias estimation plans for evaluation purposes. The handout is attached. Rita first reviewed the method that was used to produce correlation bias estimates for the 1990 PES evaluation. She then described four models which were considered in 1990 to produce the sub-national evaluation estimates and totals, and described plans for Census 2000. Demographic Analysis correlation bias estimates are produced at a national level. The challenge for evaluating the effects of correlation bias is to produce estimates for sub-national areas in order to evaluate the A.C.E. The sub-national correlation bias estimates will not be available until mid 2002 for the A.C.E. evaluations.

In 1990, we used Demographic Analysis (DA) sex ratios rather than DA totals to dampen concerns about uncertainty in DA estimates for population totals. Gregg Robinson noted in the previous ESCAP meeting that sex ratios were less subject to uncertainty. In 1990 the following approach was used by Bell, as stated in the handout, to obtain estimates of correlation bias at the national level.

> Use the DSE for females (which assumes no correlation bias) and select a model for males that produces alternative post stratum DSEs allowing for some dependence between the census and the coverage measurement survey. Then estimate the dependence by controlling the alternative male DSEs to reproduce DA sex ratios, assuming female DSEs are correct, when aggregated to the national level (Bell 1993). The national level of correlation bias was then taken to be the difference between the resulting DSE for males and the usual DSE for males. No adjustments were made for "other" sources of bias (Mulry, 1991). The correlation bias was then distributed to the PES post strata proportional to the estimate of the number of males in the fourth cell of the DSE for the post stratum (CAPE, 1992).

The fourth cell represents persons missed in both the census and the A.C.E. For post-strata with negative cell estimates, no correlation bias was assigned. The negative values resulted from the use of "unbiased" estimators of the fourth cell with high sampling variance. One of the strengths of the DSE is that it does not require direct estimation of the fourth cell.

Rita talked about alternatives we are considering for 2000 which include the four 1990 approaches and an approach developed by Elliott and Little. The four 1990 approaches included:

> The Fixed Odds Ratios Model which assumes that the odds of appearing in the A.C.E. given that an individual was enumerated in the census relative to the odds of appearing in the A.C.E. given that an individual was not enumerated in the census are arbitrary for

males and equal to 1 for females in a one-stratum design (Wolter 1990).

The Fixed Relative Risk Model which assumes a constant relative risk for enumeration in the census and A.C.E. for males and independence for females (Bell 1993).

The Fixed Sex Ratio Model which assumes that the sex ratio in the fourth cell is constant across strata (Bell 1993).

The Generalized Behavioral Response Model which assumes the probability of being included in the A.C.E. given that an individual was not enumerated in the census divided by the probability of being included in the census is constant across the post-strata for males (Bell 1993).

The Fixed Odds Ratio Model was used for the analysis in 1990. The approach developed by Elliott and Little improves the 1990 models. Using some general principles, they choose a model and embed it with a more comprehensive Bayesian model. Rita also discussed an alternative approach from Haberman, Jiang, and Spencer which is similar to that which was used in 1990 but does not assume that capture probabilities are constant for individuals. We concluded the discussion by noting that Rita and staff need to conduct additional exploration and conduct technical consultation with outside experts prior to selecting the methodology that will be used for 2000. We would consider implementing sensitivity analysis as a method of looking at how much variance is possible among the various model assumptions.

II. Targeted Extended Search - Change to Procedure

The Targeted Extended Search (TES) is an operation that is designed to reduce high variances that are caused by large-scale geocoding error in a subset of clusters. Geocoding error causes both erroneous enumerations and nonmatches on the census side and ACE side. During the TES operation, the blocks surrounding a cluster are searched for nonmatched and erroneously enumerated people.

A needed change has been discovered in the way housing units added to the initial census were treated in the TES procedure. For TES clusters, the census housing unit will be treated as a correct enumeration if it is located in a surrounding block. When needed, cases are sent to the field for map spotting. We have found problems with our TES procedure resulting from housing units added to the initial census since January. These housing units were not included in the TES operation. Our planning up to this point was based on an expectation that this would not adversely affect the A.C.E. processes or estimates. More recent analysis has indicated that we must include these housing units in our estimation process. We will do this by developing an imputation process to assign enumeration status for these units. We will impute correct enumeration status for census geocoding errors where we do not know if it is in the surrounding

block or some other block further away.

We will finalize this modification to our estimation before any data are available for review.

III. A.C.E. Quality Analysis

Howard Hogan has been directing the development of a proposal for data that will be presented to the ESCAP for its deliberation in deciding whether to use statistical correction for redistricting data. Howard presented the preliminary quality indicators that have been identified for the analysis of the initial census and the A.C.E. results. The handout is attached.

IV. Next Meeting

The next meeting scheduled for Wednesday August 9, 2000 will continue to discuss the process for evaluating the initial census and the A.C.E. results.

ESCAP Committee

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Kenneth Prewitt
Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogan
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 13 - 08/09/00

AGENDA

There was no agenda developed or used for the August 9, 2000 meeting.

ESCAP MEETING NO. 13 - 08/09/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 13

August 9, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The thirteenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 9, 2000 at 10:30. The agenda for the meeting was the A.C.E. quality indicator analysis preparation.

Committee Attendees:

William Barron
Bob Fay
John Thompson
Howard Hogan
Ruth Ann Killion
Susan Miskura

Other Attendees:

Kenneth Prewitt
Donna Kostanich
Louisa Miller
Raj Singh
David Whitford
Debbie Fenstermaker
Michael Ikeda
Maria Urrutia
Annette Quinlan

I. A.C.E. Quality Indicator Analysis

Due to staff schedules, the full ESCAP could not meet. Rather than cancel the meeting, Howard Hogan conducted an informal discussion on the progress in developing the set of A.C.E. quality indicators that will be assessed as part of the Committee's deliberations. This

discussion will continue at the next ESCAP meeting.

II. Next Meeting

The next meeting scheduled for Wednesday August 23, 2000 will discuss how 1990 data will be used to inform the 2000 decision process.

ESCAP Committee

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Kenneth Prewitt
Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogan
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 14 - 08/23/00

AGENDA

There was no agenda developed or used for the August 23, 2000 meeting.

ESCAP MEETING NO. 14 - 08/23/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 14

August 23, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The fourteenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 23, 2000 at 10:30. The agenda for the meeting was to continue the discussion on how census data will be used to inform the 2000 decision process.

Committee Attendees:

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson
Jay Waite
Bob Fay
Howard Hogan
John Long
Susan Miskura

Other Attendees:

Donna Kostanich
Raj Singh
Tommy Wright
Debbie Fenstermaker
David Hubble
Carolee Bush
Nick Birnbaum
Maria Urrutia
Annette Quinlan

I. A.C.E. Quality Analysis

The ESCAP continued the discussion of the A.C.E. measures of quality from the previous meeting. The ESCAP also discussed the presentation of these materials for the meeting sponsored by the National Academy of Sciences scheduled for October 2, 2000. Howard Hogan and Census Bureau staff will prepare reports on how we plan to analyze data for the following areas:

- Assessing Results from the Accuracy and Coverage Evaluation
- Overall Quality Indicators
- Decomposition of Dual System Estimation (DSE) Components
- Missing Data Results
- Person Matching and Follow-up Results
- Person Interviewing Results
- Dual System Estimation Results
- Demographic Analysis Results
- Variance Estimates by Size of Geographic Area
- Consistency of Post-Stratification Variables
- Quality of Census Processes
- Synthetic Assumptions
- Correlation Bias Results

We also discussed various methods by which the overall data can be synthesized. This discussion will continue at the next ESCAP meeting.

II. Next Meeting

The next meeting scheduled for Wednesday September 13, 2000 will continue the discussion on the preparation for the October 2 meeting with NAS.

ESCAP Committee

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson, Chair
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

cc:

Kenneth Prewitt
Teresa Angueira
Bill Bell
Debbie Bolton
Genny Burns
Carolee Bush
Gerald Gates
Ed Gore
Dave Hubble
Donna Kostanich
Ellen Lee
Charlene Leggieri
Don Malec
Betsy Martin
Catherine Miller
Fay Nash
Sally Obenski
Miguel Perez
Ed Pike
Magdalena Ramos
Gregg Robinson
Raj Singh
Maria Urrutia
Signe Wetrogan
David Whitford
Henry Woltman
Tommy Wright

ESCAP MEETING NO. 15 - 09/13/00

AGENDA

Annette M Quinlan
09/12/2000 08:05 AM

To: Kathleen P Zveare/DMD/HQ/BOC@BOC
cc: Maria E Urrutia/DMD/HQ/BOC@BOC
Subject: Agenda for ESCAP Meeting #15

Agenda for ESCAP Meeting #15, September 13, 2000

We will continue the discussion on preparation for October 2 meeting with NAS.

ESCAP MEETING NO. 15 - 09/13/00

HANDOUTS

DRAFT

September 8, 2000

Outline of
Assessing the Results from the
Accuracy and Coverage Evaluation

Howard Hogan

1. The purpose of this document

2. Review of A.C.E. operations
   - 2.1 Were the steps between processing and estimation properly carried out?
   - 2.2 Were the A.C.E. operations well conducted and well controlled?

3. Review of the measures of A.C.E. quality
   - 3.1 Individual Components of A.C.E. quality
     - 3.1.1 What is the level of A.C.E. Sampling?
     - 3.1.2 Consisted reporting of Census Day residence
     - 3.1.3 Matching Error
     - 3.1.4 A.C.E. Fabrications
     - 3.1.5 Missing Data
     - 3.1.6 Balancing error
     - 3.1.7 Error in measuring erroneous enumerations
     - 3.1.8 Correlation Bias
     - 3.1.9 Synthetic Bias and Synthetic Variability
     - 3.1.10 Other measurement and technical errors
       - Technical Ratio bias
       - Contamination error
       - Inconsistent Poststratification
   - 3.2 Synthesizing the components of A.C.E. Quality
     - 3.2.1 How do the individual components of A.C.E. quality combine to affect the accuracy of the population estimates?
     - 3.2.2 How does these accuracy of the A.C.E. compare to the accuracy of the uncorrected census?

4. Comparison with historical patterns and independent benchmarks
   - 4.1 Comparison with Demographic Analysis and Demographic Projections
   - 4.2 Comparison with historical patterns
   - 4.3 External measures of Census Quality
   - 4.4 Census Quality Assurance indicators
   - 4.5 Other reports of Census Quality

5. Forming an overall Assessment

Data and Analysis to Inform the ESCAP Decision

1. Introduction.

2. Review of the Quality of the Uncorrected Census.

   2.1 **Comparison with Demographic Analysis and Demographic Estimates**.
   2.2 **Direct Measures of Census Quality**.

3. Review of A.C.E. Operations.

   3.1 **Proper Execution of the Steps Between Processing and Estimation.**
   3.2 **Conduct and Control of the A.C.E. Operations.**

4. Review of Measures of A.C.E. Quality.

   4.1 **Individual Components of A.C.E. Quality**.

      4.1.1 *Sampling Variance*.
      4.1.2 *Consistent Reporting of Census Day Residence*.
      4.1.3 *Matching Error*.
      4.1.4 *A.C.E. Fabrications*.
      4.1.5 *Missing Data.*
      4.1.6 *Balancing Error.*
      4.1.7 *Erroneous Enumerations.*
      4.1.8 *Correlation Bias.*
      4.1.9 *Synthetic Bias and Synthetic Variability.*
      4.1.10 *Other Measurement and Technical Errors.*

   4.2 **Synthesizing A.C.E. Quality.**

      4.2.1 *Combining the Components of A.C.E. Quality to Assess Accuracy*.
      4.2.2 *Comparing the Accuracy of the A.C.E. to the Accuracy of the Uncorrected Census*.

5.0 Forming an Overall Assessment

ESCAP MEETING NO. 15 - 09/13/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 15

September 13, 2000

Prepared by: Nick Birnbaum.

The fifteenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 13, 2000 at 10:30. The meeting did not run the allotted hour and a half because several ESCAP members had to depart early due to other commitments.

The agenda for the meeting was to continue the discussions from previous meetings regarding drafts of the documents to be presented to the National Academy of Sciences (NAS) panel on October 2.

Committee Attendees:

Nancy Potok
Paula Schneider
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan

Other Attendees:

Kenneth Prewitt
Donna Kostanich
Raj Singh
Tommy Wright
Debbie Fenstermaker
Roxie Jones
Louisa Miller
Gregg Robinson
Nick Birnbaum
Kathleen Styles
Maria Urrutia
Annette Quinlan

I. Comparison of A.C.E. and Census Quality

Howard Hogan presented an outline of an overview document that he will draft to summarize the various analyses that the set of documents will contain. An introductory section of the overview document will define the purpose of the series of documents. Second, the analysis reports relating to the A.C.E. operations and the transitional steps will be reviewed to determine how well the specific operations and steps were executed and documented. Third, a discussion of the reports relating to measures of A.C.E. quality will be presented and include a determination of the validity of each component. Next, there will be an examination of the reports comparing the initial census data to historical patterns and independent benchmarks to determine how well the initial census was conducted. The last section will discuss the analysis report relating to the formulation of an overall assessment.

During the discussion, ESCAP members emphasized the fact that their task is to examine the quality of both the initial census counts and the A.C.E., analyze the deficiencies in both, and determine if the results of the A.C.E. could be used to improve the accuracy of the initial counts. That is, the Committee has not pre-judged the superior accuracy of the A.C.E.-based results; these results would only be applied if they can correct the deficiencies in the initial counts. Howard Hogan agreed that this point would be clearly articulated in the overview document, and that the individual analysis reports would reflect the fact that the Committee would be examining the assessments of both the initial counts and the A.C.E. results.

Attached are the outline that was discussed at the meeting and a revised version. Comments on the analysis reports were to be sent to John Thompson and Howard Hogan before c.o.b. September 20.

II. Next Meeting

The next meeting, scheduled for Wednesday September 27, 2000, will continue discussions of the documents to be presented at the NAS panel workshop.

Attachments

ESCAP MEETING NO. 16 - 09/27/00

AGENDA

Kathleen P Zveare
09/27/2000 09:01 AM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC
cc:
Subject: Agenda for Today's ESCAP Meeting

The agenda for today's ESCAP meeting from 10:30-12 in Rm. 2412/3 is as follows:

Continue preparations for the October 2 National Academy of Science meeting.

ESCAP MEETING NO. 16 - 09/27/00

HANDOUTS

Decennial Statistical Studies Division
Census 2000 Procedures and Operations Memorandum Series

Chapter B: A.C.E. Review

Chapter Code	Subject
B-1	Data and Analysis to Inform the ESCAP Recommendation
B-2	Overall Census and A.C.E. Quality Indicators
B-3	Quality of Census 2000 Processes
B-4	Accuracy and Coverage Evaluation Survey: Demographic Analysis Results
B-5	Accuracy and Coverage Evaluation Survey: Person Interviewing Results
B-6	Accuracy and Coverage Evaluation Survey: Person Matching and Followup Results
B-7	Accuracy and Coverage Evaluation Survey: Missing Data Results
B-8	Accuracy and Coverage Evaluation Survey: Decomposition of Dual System Estimate Components
B-9	Accuracy and Coverage Evaluation Survey: Dual System Estimation Results
B-10	Accuracy and Coverage Evaluation Survey: Consistency of Post-Stratification Variables
B-11	Accuracy and Coverage Evaluation Survey: Variance Estimates by Size of Geographic Area
B-12	Accuracy and Coverage Evaluation Survey: Correlation Bias

Chapter Code	Subject
B-13	Accuracy and Coverage Evaluation Survey: Comparing Accuracy
B-14	Accuracy and Coverage Evaluation Survey: Synthetic Assumptions
B-15	Accuracy and Coverage Evaluation Survey: Contributions of Service Based Enumeration Multiplicity Estimation to Corrected Census Results
B-16	Demographic Full Count Review Report

ESCAP MEETING NO. 16 - 09/27/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 16

September 27, 2000

Prepared by: Nick Birnbaum.

The sixteenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 27, 2000 at 10:30.

The agenda for the meeting was to continue the discussions from previous meetings regarding the documents to be presented to the National Academy of Sciences (NAS) Panel on October 2.

Committee Attendees:

William Barron
Nancy Potok
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan
Susan Miskura
Ruth Ann Killion

Other Attendees:

Kenneth Prewitt	Donna Kostanich
Tommy Wright	Debbie Fenstermaker
Roxie Jones	Louisa Miller
Gregg Robinson	Nick Birnbaum
Kathleen Styles	Maria Urrutia
Annette Quinlan	Carolee Bush

I. Preparations for the October 2 NAS Panel Workshop

John Thompson opened the meeting and asked Dr. Prewitt to briefly discuss his planned opening remarks for the NAS Panel meeting. Dr. Prewitt discussed how he planned to address, among other things, the rationale for not making partial or preliminary data publicly available during the ESCAP decision-making process:

- To avoid confusion
- To avoid the appearance of political manipulation
- The ESCAP needs to deliberate without external scrutiny, pressure, or influence.

The data will be made available after the decision has been made. To date, the Census Bureau, as a statistical agency, has exercised more transparency regarding its statistical programs than other statistical agencies would perhaps be comfortable with providing. There has been a great deal of pre-specification in order to support this transparency. However, this pre-specification is not always optimal. It limits changes to the methodology, in response to unanticipated circumstances, that would result in improvements to the data.

There was some discussion regarding the timing of the issuance of the Panel's report. It was noted that although we might receive some informal feedback as a result of our presentation at the October 2 workshop, the workshop would not result in formal recommendations from the Panel.

Howard Hogan then reviewed his overview document for the NAS presentation, as this would be the basis for leading the discussion at the workshop. With regard to analysis report #13, Howard stated that it was important to emphasize that the loss function analyses would not be determinative – that they would be one component of the Committee's assessment of both the initial counts and the A.C.E. results. Howard indicated that it was important to reiterate that some data that are relevant to the loss function analysis would not be available within the time frame for the Committee to make a recommendation, and would have to be estimated based on 1990 data. The specifics of the estimation were deferred for future discussion.

[Note: Howard Hogan's overview document and the rest of the final prototype analysis reports provided to the NAS Panel on October 2, 2000, are available upon request. For the convenience of the reader, an index of those documents is attached.]

In wrapping up the meeting, a couple of administrative issues were discussed. It was suggested that the points contained in the section entitled "Forming an Overall Assessment," which is the very last section of the overview document, be discussed at the beginning of the meeting, in the

event that time constraints did not enable Howard to walk through the entirety of his overview document.

It was announced that revised versions of 14 of the 16 documents would be distributed later today and that revised versions of the remaining two documents would be distributed tomorrow. Also, a meeting agenda, a document index, and directions to the National Academy would be distributed to Committee members.

II. Next Meeting

The next meeting, scheduled for October 11, 2000, will address the effect of late census data on Dual System Estimation and examine the dual system estimate variances from the 1990 PES.

Attachment

ESCAP MEETING NO. 17 - 10/11/00

AGENDA

Kathleen P Zveare
10/10/2000 10:27 AM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Elizabeth Centrella/DSD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC
cc:
Subject: Agenda for 10/11 ESCAP Meeting

The agenda for the October 11 ESCAP Meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

1. Update on Census Processes - Jay Waite

2. NAS Reports - Howard Hogan

ESCAP MEETING NO. 17 - 10/11/00

HANDOUTS

Dual System Estimate. For a given post-stratum, the formula for the dual system estimate is as follows:

$$DSE = (C - II)\left(\frac{CE}{N_e}\right)\left(\frac{N_p}{M}\right) \quad (1)$$

where

C	=	the census count;
II	=	the number of census people with insufficient information;
CE	=	the estimated number of correct enumerations from the E Sample;
N_e	=	the estimated number of people from the E Sample;
N_p	=	the estimated total population from the P Sample;
M	=	the estimated number of persons from the P-sample population who match to the Census.

Persons in Group Quarters and the Remote Alaska type of enumeration area are excluded from the 2000 A.C.E., and thus from the above numbers. For the 1990 PES, persons in the Remote Alaska type of enumeration area were excluded while persons from Group Quarters were included in these numbers.

The 2000 A.C.E. and the 1990 PES differ procedurally in their treatment of movers; that is persons whose location at the time of the survey interview differ from their location on Census Day. See the section on Movers.

Coverage Correction Factor. The coverage correction factor (CCF) is a measure of correction to assess the degree of net overcount or net undercount of the household population within the Census. The coverage correction factor (CCF) for a post-stratum is the ratio of the DSE over the census count.

$$CCF = \frac{DSE}{C} \quad (2)$$

For example, a coverage correction factor of 1.05 would imply that for every 100 people within the given post-stratum, there is a net undercount of five persons.

Table B-1: 2000 A.C.E. Results--Total Race/Hispanic Origin Domain

Total	Domain 1					
(AI on Res)	Domain 2					
(AI off Res)	Domain 3					
(Hispanic)	Domain 4					
(Black)	Domain 5					
(NH or PI)	Domain 6					
(Asian)	Domain 7					
(Wt or Oth)	Total					
Census Counts						
Data-Defined Persons (DD)		987,654,321	This number	is for display	purposes	only.
Insufficient Information (II)						
Total Persons (C)						
P Sample						
Nonmover Sample Size						
Inmover Sample Size						
Outmover Sample Size						
Weighted Nonmovers (N_n)						
Weighted Inmovers (N_i)						
Weighted Outmovers (N_o)						
Weighted Nonmover Matches (M_n)						
Weighted Outmover Matches (M_o)						
Weighted P-Sample Persons (N_p)						
Weighted P-Sample Matches (M)						
E Sample						
E-Sample Size						
Correct Enumeration Sample Size						
Weighted E-Sample Persons (N_e)						
Weighted Correct Enumerations (CE)						
Estimates						
Dual System Estimate (DSE)						
Standard Error (SE)						
Coefficient of Variation (CV) (%)						
Coverage Correction Factor						
Standard Error (SE)						
Coefficient of Variation (CV) (%)						
Net Undercount Percent (UC) (%)						
Standard Error						

Table 4. Estimates for Revised Post-Strata Groups

	Percent undercount					Standard errors				
	All	*NE*	*S*	*MW*	*W*	*All*	*NE*	*S*	*MW*	*W*
Non-Hispanic White and Other										
Owner										
Large Urbanized Areas		−2.13	.68	−.26	−.34		1.08	.71	.39	.65
Other Urban		−1.08	.52	−.10	.62		.49	.42	.40	.58
Non-Urban		−.54	.18	−.71	.29		.70	.69	1.18	.69
Nonowner										
Large Urbanized Areas		1.16	2.56	2.33	3.18		1.39	1.48	1.61	1.62
Other Urban		3.41	3.20	1.23	4.49		1.51	1.74	1.09	1.34
Non-Urban		6.52	6.23	2.85	6.08		4.20	1.71	1.51	1.81
Black										
Owner										
Large Urbanized Areas		1.63	2.16	.81	6.10		1.91	.90	.87	1.91
Other Urban	1.34					.98				
Non-Urban	3.52					1.90				
Nonowner										
Large Urbanized Areas		8.37	6.27	5.99	9.96		1.61	1.90	1.68	2.72
Other Urban	4.15					1.18				
Non-Urban	4.62					5.33				
Non-Black Hispanic										
Owner										
Large Urbanized Areas		.67	2.53	−4.33	2.89		4.45	.90	2.58	.87
Other Urban	.94					1.64				
Non-Urban	2.73					2.69				
Nonowner										
Large Urbanized Areas		6.72	9.34	6.64	5.91		3.51	2.59	3.26	1.84
Other Urban	6.60					2.74				
Non-Urban	15.80					5.01				
Asian and Pacific Islander										
Owner	−1.45					1.50				
Nonowner	6.96					2.52				
Reservation Indians	12.22					4.73				

Table 5. Counts and Undercount Rates by State: Adjustment and Revised

Name	Census count*	Adjustment estimate			Revised estimate		
		Count*	Rate	SE	Count*	Rate	SE
Alabama	4,041	4,146	2.5	.4	4,113	1.8	.3
Alaska	550	561	1.9	.4	561	2.0	.4
Arizona	3,665	3,790	3.3	.5	3,754	2.4	.5
Arkansas	2,351	2,403	2.2	.4	2,392	1.7	.3
California	29,760	30,888	3.7	.4	30,595	2.7	.4
Colorado	3,294	3,376	2.4	.5	3,364	2.1	.4
Connecticut	3,287	3,306	.6	.6	3,308	.6	.4
Delaware	666	687	3.0	.4	678	1.8	.4
DC	607	639	5.0	.5	628	3.4	.9
Florida	12,938	13,278	2.6	.4	13,197	2.0	.4
Georgia	6,478	6,633	2.3	.4	6,619	2.1	.3
Hawaii	1,108	1,136	2.5	.5	1,129	1.9	.8
Idaho	1,007	1,035	2.8	.5	1,029	2.2	.4
Illinois	11,431	11,592	1.4	.4	11,544	1.0	.4
Indiana	5,544	5,586	.7	.4	5,572	.5	.4
Iowa	2,777	2,807	1.1	.5	2,788	.4	.4
Kansas	2,478	2,506	1.2	.4	2,495	.7	.4
Kentucky	3,685	3,768	2.2	.4	3,746	1.6	.4
Louisiana	4,220	4,332	2.6	.4	4,314	2.2	.4
Maine	1,228	1,240	1.0	.6	1,237	.7	.6
Maryland	4,781	4,869	1.8	.4	4,882	2.1	.4
Massachusetts	6,016	6,039	.4	.5	6,045	.5	.5
Michigan	9,295	9,404	1.2	.4	9,361	.7	.4
Minnesota	4,375	4,419	1.0	.4	4,394	.4	.4
Mississippi	2,573	2,632	2.2	.4	2,629	2.1	.4
Missouri	5,117	5,184	1.3	.4	5,149	.6	.4
Montana	799	822	2.8	.5	818	2.4	.5
Nebraska	1,578	1,595	1.0	.4	1,589	.6	.4
Nevada	1,202	1,232	2.4	.5	1,231	2.3	.4
New Hampshire	1,109	1,116	.6	.5	1,119	.8	.5
New Jersey	7,730	7,836	1.4	.5	7,774	.6	.6
New Mexico	1,515	1,586	4.5	.5	1,563	3.1	.5
New York	17,990	18,304	1.7	.5	18,262	1.5	.6
North Carolina	6,629	6,815	2.7	.4	6,753	1.8	.3
North Dakota	639	648	1.4	.5	643	.7	.5
Ohio	10,847	10,933	.8	.4	10,922	.7	.4
Oklahoma	3,146	3,214	2.1	.4	3,203	1.8	.3
Oregon	2,842	2,898	1.9	.4	2,896	1.9	.4
Pennsylvania	11,882	11,957	.6	.5	11,917	.3	.5
Rhode Island	1,003	1,006	.3	.6	1,005	.1	.6
South Carolina	3,487	3,590	2.9	.4	3,559	2.0	.4
South Dakota	696	707	1.5	.5	703	1.0	.5
Tennessee	4,877	5,012	2.7	.4	4,964	1.7	.3
Texas	16,987	17,551	3.2	.4	17,470	2.8	.4
Utah	1,723	1,757	1.9	.5	1,753	1.7	.5
Vermont	563	571	1.4	.7	569	1.1	.8
Virginia	6,187	6,353	2.6	.4	6,314	2.0	.4
Washington	4,867	4,987	2.4	.4	4,958	1.8	.4
West Virginia	1,793	1,842	2.6	.4	1,819	1.4	.4
Wisconsin	4,892	4,924	.7	.4	4,922	.6	.4
Wyoming	454	466	2.7	.5	464	2.2	.4

* All counts in thousands.

Appendix: Table A.1. Adjustment Dual-System Estimates

Percent Undercount by Post-Stratum Group

	Direct				Smoothed			
North East	All Other	Black	Hispanic	Asian	All Other	Black	Hispanic	Asian
New England								
Central Cities	-1.74	5.69			-1.16	4.25		
Non Central City MSA	0.61				0.19			
Other Places 10,000+	0.54	5.88 *			0.59	5.39 *		
Other areas	1.68				1.79			
Middle Atlantic								
New York City CC's								
Non-owner	2.06	6.44	4.00	9.47	0.87	7.76	1.73	10.50
Owner	-2.64	-2.86			-0.23	-0.15		
Other Large MSA Central city								
Non-owner	-6.41	10.78	9.91		-0.37	7.74	2.01	
Owner	-2.93	2.66			-0.19	-0.03		
Central cities of Small MSA	2.05	17.92			0.07	9.34		
Non Central City in NYC PMSA	5.03	5.63			0.42	6.73		
Non Central City in Other Large MSA	-0.80				0.36			
Non Central City in Small MSA	-0.78	5.88 *			-0.09	5.39 *		
Other Places 10,000+	1.36				0.41			
Other areas	0.43				0.70			
South								
South Atlantic								
Large MSA Central city								
Non-owner	11.49	10.46			5.00	9.33		
Owner	1.09	1.68	2.77		1.72	0.95	4.92	
Central cities of Small MSA	2.84	4.93			2.74	4.00		
Non Central City in Large MSA	0.93	4.17	13.79		0.44	1.97	5.13	
Non Central City in Small MSA	3.50	0.27			2.80	3.59		
Other Places 10,000+	1.23	-1.71			1.51	1.60		
Other areas	3.25	5.68			2.71	2.64		
East South Central								
Large MSA Central city								
Non-owner	2.17	6.46			4.80	5.81		
Owner	3.19				2.56			
Central cities of Small MSA	0.90				2.58			
Non Central City in MSA	1.42	4.82			2.31	2.26		
Other Places 10,000+	-6.02				1.84			
Other areas	-0.95				1.65			
West South Central								
Houston,Dallas, Ft.Worth CC's								
Non-owner	6.24				4.60			
Owner	0.56	8.09	8.96		1.49	6.64	7.11	
Other Large MSA Central city								
Non-owner	1.34				3.23			
Owner	-1.16	4.54	3.18		0.69	4.82	3.76	
Central cities of Small MSA	-3.16				2.48			
Non Central City in MSA	2.07				2.28			
Other Places 10,000+	1.19	1.66	2.36		1.25	2.28	5.11	
Other areas	1.72				1.96			

	Direct				Smoothed			
	All Other	Black	Hispanic	Asian	All Other	Black	Hispanic	Asian
Midwest								
East North Central								
Chicago Detroit CC's								
Non-owner	2.76	6.76	0.38		5.17	5.77	-1.61	
Owner	-0.05	0.42			1.12	1.98		
Other Large MSA Central city								
Non-owner	1.56	4.03			1.04	4.49		
Owner	-1.24	7.09			-0.15	0.64		
Central cities of Small MSA	1.76	4.61			2.09	5.44		
Non Central City in Large MSA	0.84				0.59			
Non Central City in Small MSA	0.96	3.99 *			0.64	4.66 *		
Other Places 10,000+	0.42				0.20			
Other areas	-1.64				-0.99			
West North Central								
Large MSA Central city								
Non-owner	5.20	5.47			2.47	5.44		
Owner	-0.53				-0.33			
Central cities of Small MSA	1.82	4.85			1.90	7.23		
Non Central City inLarge MSA	1.09				0.71			
Non Central City in Small MSA	0.22	3.99 *			1.64	4.66 *		
Other Places 10,000+	0.83				0.75			
Other areas	0.78				0.31			
West								
Mountain								
Large MSA Central city								
Non-owner	4.65	1.48			5.03	4.61		
Owner	1.24				0.98			
Central cities of Small MSA	2.88				1.52			
Non Central City in MSA	0.60	7.39 *			0.75	7.80 *		
Other Places 10,000+	1.22				1.45			
Other areas	3.00				3.22			
Pacific								
Non-owner								
Los Angeles/Long Beach CC's	6.44	7.38	10.14	6.29	4.75	6.83	7.87	6.50
Other Large MSA Central city	3.73				3.72			
Central cities of Small MSA								
Owner								
Los Angeles/Long Beach CC's	-0.35	8.36	2.01	3.10	1.39	7.86	1.95	4.80
Other Large MSA Central city	1.39				1.39			
Central cities of Small MSA								
Central cities of Small MSA	0.56				0.95			
Non Central City in Large MSA/PMSA	1.05	14.32	5.65	0.82	0.17	16.37	6.94	0.79
Non Central City in Small MSA	2.90				3.15			
Other Places 10,000+	1.38	7.39 *		-3.22	1.89	7.80 *		0.18
Other areas	3.15				1.92			
Reservation Indian				12.72				12.72

NOTE: Bold indicates cell is significantly different from zero at 90% level. Boxes show which cells were combined to form post-strata. Asians are included in All Other when not separately shown. * indicates that the cell is combined with another non-adjacent cell.

[*Received January 1992. Revised November 1992.*]

REFERENCES

Belin, T., and Diffendal, G. (1991), "Results From the Handling of Unresolved Enumeration Status, Missing Characteristic Data, and Noninterviews in the 1990 Post-Enumeration Survey," STSD Decennial Census Memorandum Series V-112, *U.S. Bureau of the Census*, Washington, DC.

Furnival, G. M., and Wilson, R. W. (1974), "Regression by Leaps and Bounds," *Technometrics*, 16, 499–512.

ESCAP MEETING NO. 17 - 10/11/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 17

October 11, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The seventeenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on October 11, 2000 at 10:30. Howard Hogan discussed the potential effects of late census data on the Dual System Estimation (DSE) process and the DSE variances from 1990.

Committee Attendees:

William Barron
Nancy Potok
Paula Schneider
Nancy Gordon
John Thompson
Jay Waite
Bob Fay
Howard Hogan
John Long
Susan Miskura

Other Attendees:

Kenneth Prewitt
Donna Kostanich
Raj Singh
Tommy Wright
Sally Obenski
Kathleen Styles
Maria Urrutia
Annette Quinlan

I. Census Processes

Jay Waite updated the ESCAP on the status of the Hundred percent Census Unedited File (HCUF) creation. At the time of the meeting, over half of the HCUF state files had been approved.

In the most recent Executive State Of the Census (ESOC) report there was a reference to the possibility of duplicated addresses in the census. Jay briefly described this situation and summarized what actions are being taken to address this issue.

II. Dual System Estimation

Howard Hogan discussed the potential effect of late census data on the Dual System Estimates (DSEs). Late census data are treated, during matching and estimation, in the same manner as any whole person substitutions in the census. Since it would be impossible to match such people, late census data and whole person substitutions are not included in the matching operations or the calculation of the DSE. However, the late census data are included with the final census counts used in calculating the Coverage Correction Factors (CCFs); that is, they are included since the CCF is the ratio, for a given post-stratum, of the DSE to the final census count, including late census data.

Howard then discussed the table which will be prepared to summarize the A.C.E. results for ESCAP deliberation purposes (attached). The table will now include an additional variable to distinguish between the census counts before late census data and after late census data.

III. 1990 Estimates and Variances

Howard Hogan discussed the direct and smoothed estimation results and variances from the 1990 Post Enumeration Survey. This was to provide background information to the Committee for them to develop an understanding of what occurred in 1990.

In 1990, two models were used for estimating the total population, a direct model and a smoothed model. The smoothed model has been eliminated for the 2000 A.C.E. The direct and smoothed percent undercount rates for 1990 are summarized in the attached tables. The smoothing made some of the estimates more reasonable and resulted in less variation.

In 1990, when the decision was made to drop plans for smoothing the estimates, it was also decided to decrease the number of post-strata and thereby increase the number of people in each post-stratum so as to reduce the variance of the estimates. This history led to a discussion

of the relative size of each post-stratum. It was decided that a useful tool in aiding the ESCAP deliberation process would be a supplementary table showing the sizes of the post-strata. This would show how large the groups were and the potential impact on the estimated coverage errors that will be calculated for the post-strata.

The ESCAP also discussed the decisions that would be addressed. It was decided that the ESCAP would focus initially on the decision regarding adjustment of the redistricting data. The ESCAP would then determine whether additional decisions are necessary.

IV. Next Meeting

The next meeting scheduled for Wednesday October 25, 2000 will discuss loss functions.

Attachments

ESCAP MEETING NO. 18 - 10/25/00

AGENDA

There was no agenda developed or used for the October 25, 2000 meeting.

ESCAP MEETING NO. 18 - 10/25/00

HANDOUTS

Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

Loss Functions - The Fundamentals

Assumes True Population Known

Table 1: "The Nation"

Post Stratum	True Population	Census Count	Census Error	DSE Estimate	DSE Bias	DSE SE	DSE Error
Owner	16,500	16,335	-165	16,246.9			-253.1
Renter	1,500	1,425	-75	1,456.2			-43.8
Total	18,000	17,760	-240	17,703.1			-296.9

Table 2: "The States"

ST	Post Stratum	True Population	Census Count	Census Error	ACE Estimate	ACE Bias	ACE SE	ACE Error
1	Owner	11,400	11,286	-114	11,225.2			-174.8
	Renter	300	285	-15	291.2			-8.8
2	Owner	500	495	-5	492.3			-7.7
	Renter	500	475	-25	485.4			-14.6
3	Owner	200	198	-2	196.9			-3.1
	Renter	100	95	-5	97.1			-2.9
4	Owner	1,300	1,287	-13	1,280.1			-19.9
	Renter	100	95	-5	97.1			-2.9
5	Owner	3,100	3,069	-31	3,052.5			-47.5
	Renter	500	475	-25	485.4			-14.6
	Total	18,000	17,760	-240	17,703.1			-296.9

Loss Functions - The Fundamentals (cont.)

Assumes True Population Known

Table 3a: Numeric Loss Function - Squared Error

State	True Population	Census Count	Census Error	Census Error^2	ACE Estimate	ACE Bias	ACE SE	ACE Error	ACE Error^2
1	11,700	11,571	-129	16,641.0	11,516			-183.6	33,713.5
2	1,000	970	-30	900.0	978			-22.3	496.2
3	300	293	-7	49.0	294			-6.0	35.9
4	1,400	1,382	-18	324.0	1,377			-22.9	522.6
5	3,600	3,544	-56	3,136.0	3,538			-62.2	3,863.1
Total	18,000	17,760	-240	21,050.0	17,703.1			-296.9	38,631.2

Census Loss 21,050.0 ACE Loss = 38,631.2

Net gain from adjustment = -17581.2

Table 3b: Distributive Loss Function - Squared Error

State	True Share	Census Share	Census Error	Census Error^2	ACE Share	ACE Bias	ACE SE	ACE Error	ACE Error^2
1	65.00%	65.15%	0.15%	0.0002%	65.05%			0.05%	0.0000%
2	5.56%	5.46%	-0.09%	0.0001%	5.52%			-0.03%	0.0000%
3	1.67%	1.65%	-0.02%	0.0000%	1.66%			-0.01%	0.0000%
4	7.78%	7.78%	0.00%	0.0000%	7.78%			0.00%	0.0000%
5	20.00%	19.95%	-0.05%	0.0000%	19.98%			-0.02%	0.0000%
Total	100.00%	100.00%	0.00%	0.0003%	100.00%			0.00%	0.0000%

Census Loss 0.0003% ACE Loss = 0.0000%

Net gain from adjustment = 0.0003%

Loss Functions - The Fundamentals (cont.)

Assumes True Population Known

Table 4a: Numeric Loss Function - Squared Error Relative to True Population

State	True Population	Census Count	Census Error	Census Error^2 / Pop.	ACE Estimate	ACE Bias	ACE SE	ACE Error	ACE Error^2 / Pop.
1	11,700	11,571	-129	1.42	11,516			-183.6	2.88
2	1,000	970	-30	0.90	978			-22.3	0.50
3	300	293	-7	0.16	294			-6.0	0.12
4	1,400	1,382	-18	0.23	1,377			-22.9	0.37
5	3,600	3,544	-56	0.87	3,538			-62.2	1.07
Total	18,000	17,760	-240	3.59	17,703.1			-296.9	4.94

Census Loss 3.59 ACE Loss = 4.94

Net gain from adjustment = -1.36

Table 4b: Distributive Loss Function - Squared Error Relative to True Share

State	True Share	Census Share	Census Error	Census Error^2 / True	ACE Share	ACE Bias	ACE SE	ACE Error	ACE Error^2 / True
1	65.00%	65.15%	0.15%	0.0004%	65.05%			0.05%	0.0000%
2	5.56%	5.46%	-0.09%	0.0016%	5.52%			-0.03%	0.0002%
3	1.67%	1.65%	-0.02%	0.0002%	1.66%			-0.01%	0.0000%
4	7.78%	7.78%	0.00%	0.0000%	7.78%			0.00%	0.0000%
5	20.00%	19.95%	-0.05%	0.0001%	19.98%			-0.02%	0.0000%
Total	100.00%	100.00%	0.00%	0.0022%	100.00%			0.00%	0.0003%

Census Loss 0.0022% ACE Loss = 0.0003%

Net gain from adjustment = 0.0019%

Loss Functions - The Fundamentals (cont.)

Assumes True Population Known

Table 5a: Numeric Loss Function - Squared Error Relative to Population Squared

State	True Population	Census Count	Census Error	Census Error^2 / Pop.^2	ACE Estimate	ACE Bias	ACE SE	ACE Error	ACE Error^2 / Pop.^2
1	11,700	11,571	-129	0.0001	11,516			-183.6	0.0002
2	1,000	970	-30	0.0009	978			-22.3	0.0005
3	300	293	-7	0.0005	294			-6.0	0.0004
4	1,400	1,382	-18	0.0002	1,377			-22.9	0.0003
5	3,600	3,544	-56	0.0002	3,538			-62.2	0.0003
Total	18,000	17,760	-240	0.0020	17,703.1			-296.9	0.0017

Census Loss 0.0020 ACE Loss = 0.0017

Net gain from adjustment = 0.0003

Table 5b: Distributive Loss Function - Squared Error Relative to Squared True Share

State	True Share	Census Share	Census Error	Census Error^2 / True^2	ACE Share	ACE Bias	ACE SE	ACE Error	ACE Error^2 / True^2
1	65.00%	65.15%	0.15%	0.0005%	65.05%			0.05%	0.0001%
2	5.56%	5.46%	-0.09%	0.0285%	5.52%			-0.03%	0.0035%
3	1.67%	1.65%	-0.02%	0.0103%	1.66%			-0.01%	0.0012%
4	7.78%	7.78%	0.00%	0.0000%	7.78%			0.00%	0.0000%
5	20.00%	19.95%	-0.05%	0.0005%	19.98%			-0.02%	0.0001%
Total	100.00%	100.00%	0.00%	0.0399%	100.00%			0.00%	0.0048%

Census Loss 0.0399% ACE Loss = 0.0048%

Net gain from adjustment = 0.0351%

ESCAP MEETING NO. 18 - 10/25/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 18

October 25, 2000

Prepared by: Nick Birnbaum.

The eighteenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on October 25, 2000 at 10:30. The agenda for the meeting was to familiarize the Committee members with the purpose of loss functions and the complexities inherent in this type of analysis. Refer to the DSSD Census 2000 Procedures and Operations Memorandum Series B-13 for a detailed discussion of the methodology to be used for conducting these analyses.

Committee Attendees:

William Barron
Nancy Potok
Nancy Gordon
Paula Schneider
Cynthia Clark
John Thompson
Ruth Ann Killion
Bob Fay
Howard Hogan
Susan Miskura
John Long

Other Attendees:

Kenneth Prewitt
Donna Kostanich
Raj Singh
Tommy Wright
Nick Birnbaum
Kathleen Styles
Maria Urrutia
Annette Quinlan
Carolee Bush

I. Brief Introduction to Loss Function Analysis

John Thompson began the meeting with a brief discussion of loss functions, summarizing what they are, how they are used, and providing an historical context (that is, the role they played in the 1990 adjustment decision process).

As has been mentioned earlier, the Census Bureau will base its decision on a comprehensive set of data and analyses -- the loss function analysis is one tool among many the Committee will utilize for examining the accuracy of the initial counts and the A.C.E. results.

At this point, John turned the meeting over to Howard Hogan, who provided the Committee with a brief A.C.E. operational update before giving his presentation on loss function analysis.

II. Loss Functions - The Fundamentals

Howard explained that loss functions are used to compare two sets of counts or share distributions -- unadjusted versus adjusted -- to determine which set is closer to the "true" count or share distribution. That is, loss functions involve comparing the census errors to the coverage measurement survey errors to determine which has the smaller "loss" when compared to the "true" counts or shares. Since the "true" count or share distribution can never be known, one has to rely on an estimated truth (a target number or share distribution) to perform the loss function analysis. Estimates of the "true" population and "true" population shares (or proportions) are produced for states and sub-state areas, depending on the desired level of analysis. These estimated "truths" have variances and biases associated with them, making the loss function analysis more complex. John briefly discussed how this comparison problem was addressed in 1990.

The input to the loss function analysis is based on a total error model used to estimate the net effect of sampling and non-sampling error in the initial census and the A.C.E. The components of the total error model are derived from the Census Bureau's evaluation studies providing various measures of sampling and non-sampling error. Because the Census Bureau will not complete some of its evaluations until late 2001 or 2002, complete information on the components of the total error model will not be available within the time frame for producing the ESCAP recommendation. Consequently, the comparison of accuracy between the adjusted and unadjusted Census 2000 population data will be modeled from 1990 census components. The methodology for developing the components of the Census 2000 total error model will be discussed in more detail at future ESCAP meetings.

III. Next Meeting

The agenda for the next meeting, scheduled for November 8, 2000, is to examine preliminary demographic analysis estimates.

ESCAP MEETING NO. 19 - 11/08/00

AGENDA

Kathleen P Zveare
11/02/2000 01:51 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC
cc:
Subject: Agenda for 11/8 ESCAP Meeting

The agenda for the November 8 ESCAP Meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

Demographic Analysis - Greg Robinson

ESCAP MEETING NO. 19 - 11/08/00

HANDOUTS

Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

Presentation on Demographic Analysis

by

J. Gregory Robinson

November 8, 2000

What is Demographic Analysis?

* Demographic Analysis (DA) represents a macro-level approach to measuring net undercount, where analytic estimates are constructed based on various types of demographic data essentially independent of the census, such as administrative statistics on births, deaths, and immigration; estimates of emigration and undocumented migration, and Medicare data. The difference between the DA estimated population and the census count provides an estimate of net census undercount.

* The demographic accounting equation:

1. Population (<65)= Births (since 1935)
   - Deaths (to persons born after 1935)
   +Immigrants (born after 1935)
   - Emigrants (born after 1935)

2. Population (65+)= Medicare Count
   + Estimated unenrolled

Table 1: Illustrative Values of DA Components for the Estimated U.S. Resident Population, April 1, 2000 (Numbers in Millions)

Component	Age in 2000				
	All Ages	Under 15	15-44	45-64	65+
Total	279.9	60.7	122.6	61.6	35.0
Under age 65:					
Births	235.6	59.6	112.3	63.7	-
Deaths	-15.3	-0.7	-5.0	-9.6	-
Immigrants	29.5	2.1	17.6	9.8	-
Emigrants	-4.9	-0.3	-2.3	-2.3	-
Ages 65+:					
Medicare +	35.0	-	-	-	35.0

Components of Change for the U.S. Resident Population, April 1, 1990 to April 1, 2000
(Numbers in thousands)

Component		Estimate
Estimated Resident Population, 4-1-1990		253,394
Components of Change, 4-1-1990 to 4-1-2000		
Births (adjusted for underregistration)	+	40,077
Deaths	-	22,721
Immigration components		
Legal immigrants (includes refugees)	+	8,381
Net migration from Puerto Rico	+	126
Net migration of temporary residents	+	7
Net migration of Federal U.S. citizens	+	312
Net movement of Armed Forces overseas	-	(237)
Net undocumented migration	+	2,765
Emigration components		
Foreign-born emigration	-	2,163
Native emigration	-	480
Estimated Resident Population, 4-1-2000		279,935

Births

* largest component by far (235.6 million from 1935-2000)
* administrative data set (from National Center for Health Statisitcs)
* estimation: to adjust for birth underregistration
  - use results of 3 tests of registration completeness)

Year	Total	White	Black
1940	92.5	94.0	81.9
1950	97.9	98.6	93.7
1964-68	99.2	99.4	98.0

- the 1940 and 1950 results are based on matching of birth certificates filed in each State with "birth cards" for infants under 3 or 4 months of age filled out by enumerators conducting the census. The 1964-68 test involved match of birth certificates filed in 1964-68 with "birth cards" for children under 5 years of age filled out by interviewers during a series of samples of the Current Population Survey (CPS) and National Health Survey between June 1969 and March 1970.
- completeness factors for years between test are interpolated; factors for years since 1968 are extrapolated from 1964-98 test results

* uncertainty on births (low and high mulipliers from assessment of 1990 DA error)

	Total		Black	
	Low	High	Low	High
1940-45	.975	1.025	.918	1.059
1965-70	.993	1.012	.989	1.018
1980-90	.992	1.020	.986	1.030

- components of error are represented by correlation bias (unregistered infants may be less likely to be enumerated in the census), match bias (registered infants may not be "matched" to census birth card because of name differences or other reasons), extrapolation error (after 1968), and sampling error (1964-68 results).
- "true" number of births are more likely to be higher than lower
- uncertainty is the same for males and females, so birth errors have little effect on sex ratios

Deaths

* smaller component, since national DA estimates for 2000 involve deaths only for the population under 65 (Medicare is estimate source for 65+). Deaths to the population under age 65 in 2000 total 15.3 million over 1935-2000 (compared to 235.6 million births).
* administrative data set (from National Center for Health Statistics)
* no national test of death underregistration
* estimation: adjustment for underregistration of infant deaths (up to 1960). Assumed that infant underregistration is one-half as large as birth underregistration; we will revisit this assumption.

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total		Black	
	Low	High	Low	High
1985-90	.995	1.008	.995	1.012

Legal Immigrants

* includes persons admitted for permanent residence (7.3 million from 1990-2000) and refugees (1.0 million from 1990-2000)
* administrative data set (from Immigration and Naturalization Service and Office of Refugee Resettlement)
* estimation: race distribution of immigrants based of race distribution specific to country of origin as reported in most recent census

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total	
	Low	High
1985-90	.995	1.15

Net Migration from Puerto Rico

* small component (0.1 million estimated net migrants for 1990-2000)
* component is estimated, using cohort component techniques and successive census data. Example:

  Expected Born in P.R. 1990 (age x) = 1980 census Born in P.R. (age x-10) * Survival rate
  Net Migration = 1990 Census Born in P.R. (age x) - Expected Born in P.R. (age x)

* Puerto-Rican migration flows for 1990-2000 are based on projecting migration levels measured for 1980-1990.

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total	
	Low	High
1980-90	.75	1.35

Net Migration of Temporary Residents

* component is assumed to be a constant stock; takes into account unique age distribution of temporary migrants (example–foreign students)
* component is estimated; based on applying algorithms to 1990 census data specific to citizenship, year of entry, age, school enrollment, occupation. Produces estimate of about 500,000 "temporary" residents
* component was first included in 1990 DA estimates; based on estimates of resident foreign students

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total	
	Low	High
1980-90	.95	1.50

Net Migration of Civilian Citizens

* small component (0.3 million estimated net migrants for 1990-2000)
* component is based in part on administrative data from Dept. of Defense and Office of Personnel Management; dependents are estimated.
* only civilian citizens affiliated with Federal Government (and dependents) are included.

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total	
	Low	High
1980-90	.50	1.50

Net Movement of Armed Forces Overseas

* usually small component, though sized varies by sex and age (net movement into U.S. of 0.2 million for 1990-2000)
* administrative data set (from Dept. of Defense)

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total	
	Low	High
1980-90	.95	1.05

Net Undocumented Migration

* most problematic component (2.8 million estimated net undocumented migrants for 1990-2000)
* component is estimated, using cohort component techniques, data from INS, census data and CPS data

    Legally resident Foreign-Born =INS admin. data (e.g., entered 1985-90)
    Total enumerated Foreign-Born=Census or CPS estimate
    Estimated undocumented (enumerated)=Enumerated F.B. - Legally resident F.B.

* component was first included in 1980 DA estimates; based on analysis of estimates of undocumenteds included in 1980 census (2 million in census, allowance for 1 million not counted produced estimate of 3 million total undocumented residents in 1980).
* component for 1990 DA estimate based on analysis of estimates of undocumenteds included in CPS surveys during 1980's (2.4 million in CPS, allowance for those not counted produced estimate of 3.3 million total undocumented residents in 1990).
* estimates of undocumented for 1990 reflect legalization of over 1.6 million persons through Immigration Reform and Control Act of 1986 (IRCA). The IRCA program provided only "administrative" data on this elusive population to benchmark against the analytic estimates. Also provided age-sex-race distribution.
* estimate of undocumented flow for 1990-2000 (275,000 annually) based on analysis of 1990 census and research with CPS nativity data.

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total	
	Low	High
1980-90	.50	1.65

Emigration

* the second problematic component of appreciable size (2.7 million estimated emigrants for 1990-2000, including 2.2 million foreign-born and 0.5 million native-born)
* component is estimated. Foreign-born emigration is estimated using cohort component techniques applied to successive census data on the foreign-born.

  Expected F.B.1990 (age x) = 1980 census F.B. (age x-10) * Survival rate + Immigrants
  Net Migration = 1990 Census F.B. (age x) - Expected F.B. (age x)

* Native-born emigration is estimated using cohort component techniques applied to census data on U.S.-born in other countries, supplemented by adminstrative data.

* Foreign-born emigration flows for 1990-2000 are based on projecting rates of emigration measured for 1980-1990. Native-born emigration flows for 1990-2000 are based on projecting levels measured circa 1980.

* Multiplicity-based estimates of emigration using special supplements to the CPS during 1980's provided rough numbers to assess the analytic estimates.

* uncertainty (low and high mulipliers from assessment of 1990 DA error)

	Total	
	Low	High
1980-90	.625	1.625

Medicare

* component based on Medicare enrollments, adjusted for the un-enrolled (preliminary estimate of 35 million for 2000).
* administrative data set (from Health Care Financing Administration)
* estimation: to adjust for persons not enrolled in Medicare. Underenrollement factors for 1990 DA based on three sources: CPS and SIPP survey estimates of percent Medicare coverage, and estimates based on cohort component analysis of time series of Medicare enrollments. Overall underenrollment is estimated to be 3.5 percent in 1990. Estimates are being developed for 2000 Medicare underenrollment.

History of DA has been:

- ★ assembling administration data to measure the components (births, deaths, immigration)
- ★ adjusting for imperfections in the component data (example-birth registration completeness)
- ★ making estimates for components with no administrative data (example-emigration and undocumented immigration)
- ★ constantly evaluation the time series of DA coverage estimates and underlying component data (for internal and longitudinal consistency)
- ★ over 50 years of experience at this

Figure 1. Percent Net Undercount by Race 1940-1990

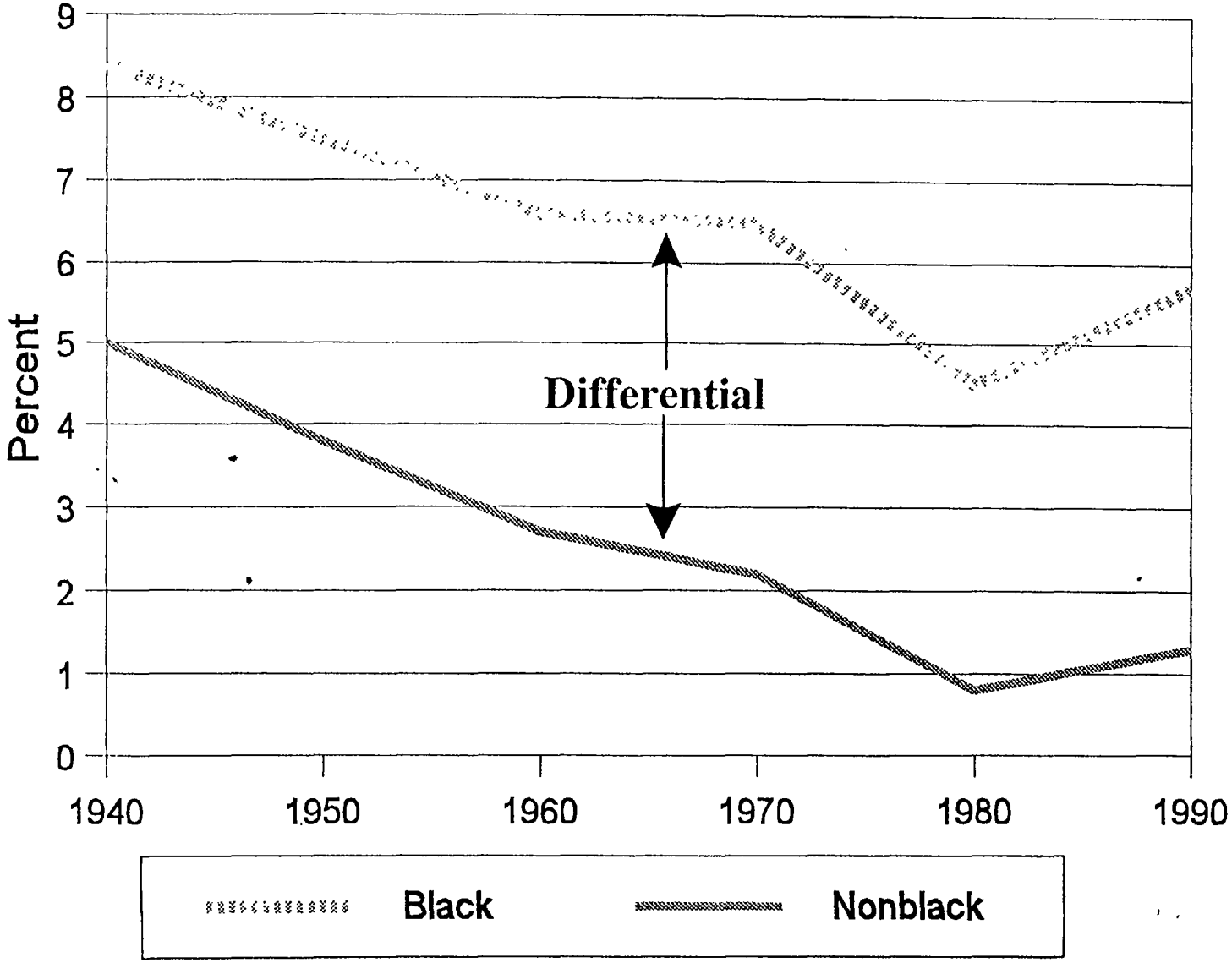
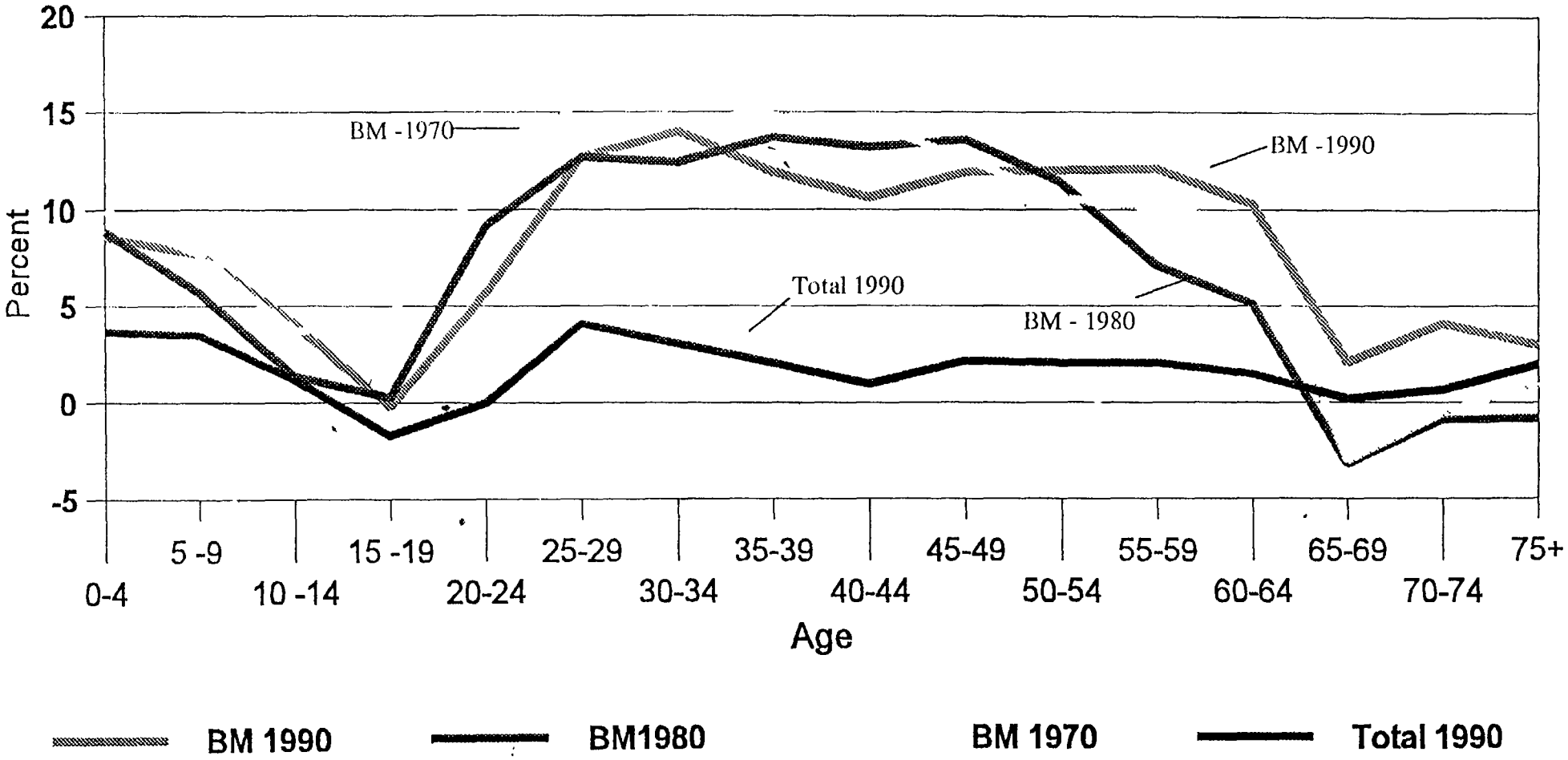


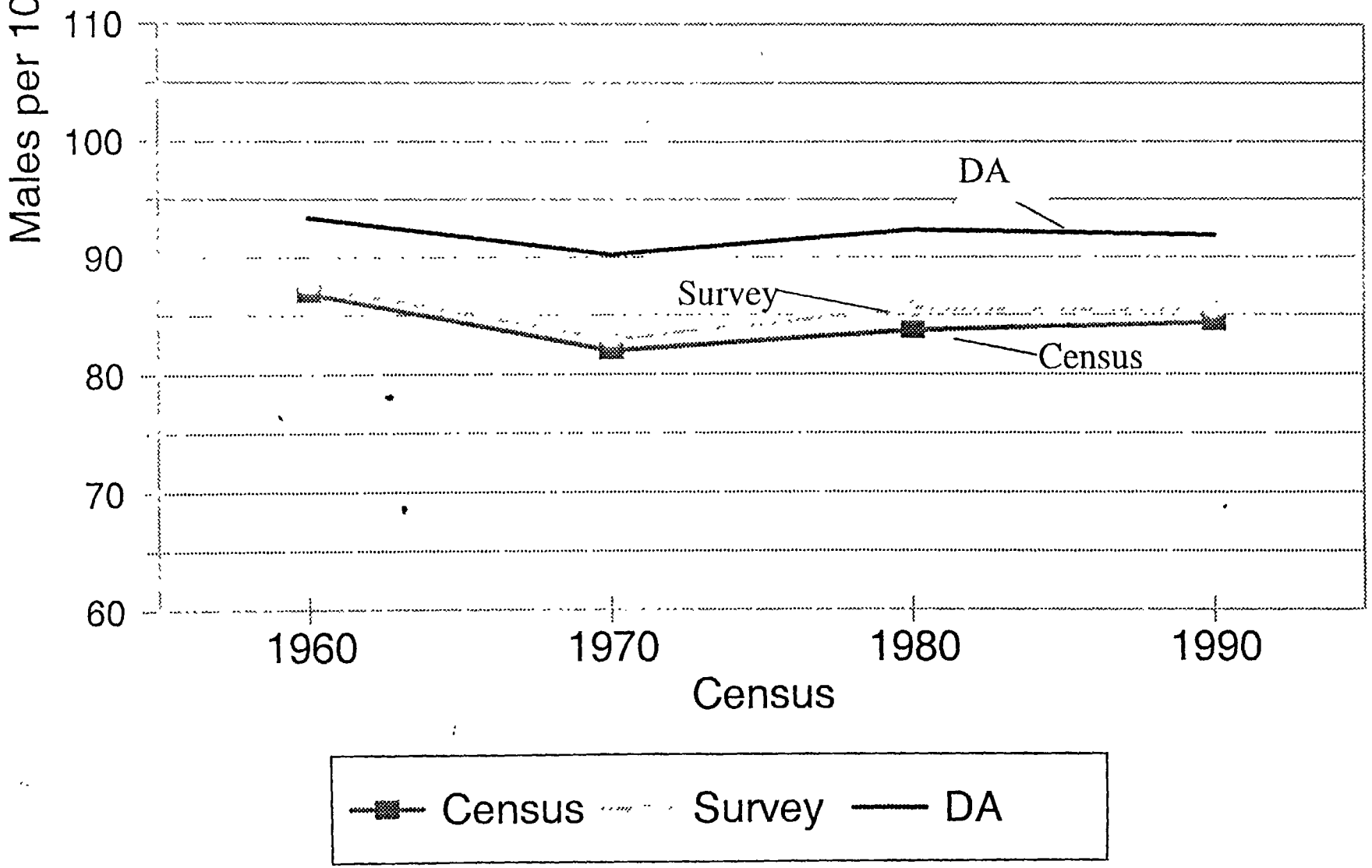
Figure 2. Percent Net Undercount: Black Males 1970 - 1990, and Total, 1990



1

1

Sex Ratios for Black Adults Census, Surveys, and DA: 1960-1990



Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

Table 1. Distribution of Births According to Alternative Race Assignment Rules: Minority Rule, Race of Father Rule, Race of Mother Rule

Year	NONBLACK							BLACK						
	Race assignment rule			Difference from Father rule		Pct Differnce		Race assignment rule			Difference from Father rule		Pct Differnce	
	Minority	Father	Mother	Minority	Mother	Minority	Mother	Minority	Father	Mother	Minority	Mother	Minority	Mother
1989	3,331,409	3,342,511	3,367,834	-11,102	25,323	-0.33	0.76	709,549	698,447	673,124	11,102	-25,323	1.59	-3.63
1990	3,433,481	3,445,299	3,473,876	-11,818	28,577	-0.34	0.83	724,731	712,913	684,336	11,818	-28,577	1.66	-4.01
1991	3,385,556	3,397,923	3,428,305	-12,367	30,382	-0.36	0.89	725,351	712,984	682,602	12,367	-30,382	1.73	-4.26
1992	3,345,420	3,358,514	3,391,381	-13,094	32,867	-0.39	0.98	719,594	706,500	673,633	13,094	-32,867	1.85	-4.65
1993	3,292,213	3,305,575	3,341,365	-13,362	35,790	-0.40	1.08	708,027	694,665	658,875	13,362	-35,790	1.92	-5.15
1994	3,263,495	3,277,386	3,316,376	-13,891	38,990	-0.42	1.19	689,272	675,381	636,391	13,891	-38,990	2.06	-5.77
1995	3,241,034	3,255,997	3,296,450	-14,963	40,453	-0.46	1.24	658,555	643,592	603,139	14,963	-40,453	2.32	-6.29
1996	3,237,605	3,253,145	3,296,713	-15,540	43,568	-0.48	1.34	653,889	638,349	594,781	15,540	-43,568	2.43	-6.83
1997	3,218,102	3,234,483	3,280,981	-16,381	46,498	-0.51	1.44	662,792	646,411	599,913	16,381	-46,498	2.53	-7.19

ESCAP MEETING NO. 19 - 11/08/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 19

November 8, 2000

Prepared by: Nick Birnbaum.

The nineteenth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on November 8, 2000 at 10:30. The agenda for the meeting was to familiarize the Committee members with the techniques of demographic analysis (DA) and to provide preliminary data on the DA estimated resident population as of April 1, 2000.

Committee Attendees:

William Barron
Nancy Potok
Paula Schneider
John Thompson
Jay Waite
Ruth Ann Killion
Howard Hogan
Susan Miskura

Other Attendees:

Raj Singh
Tommy Wright
Gregg Robinson
Signe Wetrogan
Roxie Jones
Nick Birnbaum
Kathleen Styles
Maria Urrutia
Annette Quinlan
Carolee Bush

I. Demographic Analysis -- What It Is And How It Is Used

Gregg Robinson began his presentation by explaining demographic analysis (DA). The purpose of the meeting was to explain DA to the Committee members in preparation for their analysis of the Census 2000 data.

DA represents a macro-level approach to measuring net undercount, where analytic estimates are constructed based on various types of demographic data essentially independent of the census, such as administrative statistics on births, deaths, and immigration; estimates of emigration and undocumented migration, and Medicare data. The difference between the DA estimated population and the census count provides an estimate of net census undercount or overcount.

DA uses the following demographic accounting equation:

For the population under 65:

> Births (since 1935) minus Deaths (to persons born after 1935) plus Immigrants (born after 1935) minus Emigrants (born after 1935).

For the population 65 and over:

> Medicare Count plus
> Estimated unenrolled.

The immigration and emigration components are developed from the following sub-components:

<u>Immigration components</u>:

> Legal immigrants
> Net migration from Puerto Rico
> Net migration of temporary residents
> Net migration of Federal civilian employees
> Net movement of Armed Forces overseas
> Net undocumented migration

<u>Emigration components</u>:

> Foreign-born emigration
> Native emigration

The techniques of DA allow for refinement of the estimates: for example, adjustments for imperfections in the component data, such as those relating to birth registration completeness, or the constant evaluation of time series of DA coverage estimates and the underlying component data (for internal and longitudinal consistency).

Gregg presented a preliminary DA estimate for the U.S. Resident Population as of April 1, 2000, by DA component and broad age group. He also presented data on the change in components from April 1, 1990 to April 1, 2000. Gregg then walked the Committee through the preliminary component estimates for April 1, 2000, explaining how they were developed and the uncertainties associated with them. The official April 1, 2000 DA population estimates will be available to the Committee in sufficient time to allow for review and analysis of these data to properly inform its recommendation to the Director.

Gregg also presented historical data on DA estimates of net undercount by race and age, and DA estimates of sex ratios compared with the census and coverage measurement surveys.

Finally, there was some discussion of the assignment of race for birth data in which one parent is Black and the other Non-Black, and its impact on demographic analysis estimates of the Black and Non-Black populations.

II. Next Meeting

The agenda for the next meeting, scheduled for November 22, 2000, is to examine results from the A.C.E. Before Follow-up Person Matching.

ESCAP MEETING NO. 20 - 11/22/00

AGENDA

Kathleen P Porter
11/20/2000 03:09 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Jane F Green/DSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC
cc: Kirsten K West/POP/HQ/BOC@BOC, Danny R Childers/DSSD/HQ/BOC@BOC
Subject: Agenda for Nov. 22 ESCAP Meeting

The agenda for the November 22 ESCAP Meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

1. How the Bureau will release results of demographic analysis - John Long

2. Person duplication and A.C.E. non-interview rates - Dan Childers

3. Demographic benchmark analysis of housing units - Kirsten West

ESCAP MEETING NO. 20 - 11/22/00

HANDOUTS

Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

November 27, 2000

MEMORANDUM FOR Howard Hogan
Chief, Decennial Statistical Studies Division

From: Danny R. Childers
Roxanne Feldpausch
Xijian Liu
John A. Jones
Decennial Statistical Studies Division

Subject: Accuracy and Coverage Evaluation: Results from Before Follow-up Person Matching

Attached are results from the before follow-up person matching.

1. Before Follow-up Results for Noninterviews

In 1990, 1.6 percent of the occupied PES housing units were noninterviews. Procedure B, which searches for the current residents at their census day address, was used in 1990. The percent noninterview is for the current residents which include the nonmovers and inmovers.

Table 1.1: 1990 P-Sample Housing Units

	Total Housing Units	Occupied Housing Units
Interviews	85.2	98.4
Household Member	81.2	93.7
Non Household Member	4.1	4.7
Noninterviews	1.4	1.6
Occupied Housing Units	86.6	100.0
Vacant	13.4	
Total Housing Units	100.0	

In 2000, 0.2 percent of the occupied housing units for current residents were noninterviews and 1.0 percent for census day residents. The current residents are nonmovers and inmovers and the census day residents are nonmovers and outmovers. Procedure C, which searches for the census day residents within the sample cluster, is used for 2000.

Table 1.2: 2000 Housing Units for Current Residents		
	Total Housing Units	Occupied Housing Units
Interviews	89.9	99.8
Household Member	85.1	94.5
Non Household Member	4.8	5.3
Noninterviews	0.1	0.2
Occupied Housing Units	90.0	100.0
Vacant	10.0	
Total Housing Units	100.0	

Table 1.3: 2000 Housing Units for Census Day Residents		
	Total Housing Units	Occupied Housing Units
Interviews	90.9	99.0
Household Member	84.5	92.1
Non Household Member	6.4	6.9
Noninterviews	0.9	1.0
Occupied Housing Units	91.8	100.0
Vacant	8.2	
Total Housing Units	100.0	

After the data is received in headquarters, preliminary outcome codes are assigned to the census day residents before processing begins. When the people said they lived in group quarters or had another residence on census day, they are removed from the P-sample. When the whole household was not a

resident of the housing unit, the housing unit is converted to vacant. In addition, whole households of census day residents with insufficient information for matching and follow-up are converted to noninterviews.

Table 1.4: Preliminary Outcome Code for Housing Units of Census Day Residents in 2000

	Total Housing Units	Occupied Housing Units
Interviews	88.6	98.1
Household Member	82.3	91.1
Non Household Member	5.4	6.0
Partial	0.9	1.0
Noninterviews	1.7	1.9
Occupied Housing Units	90.3	100.0
Vacant	9.7	
Total Housing Units	100.0	

The P-sample housing units for whole households of P-sample people clerically coded during the before follow-up matching as duplicates and insufficient information for matching and follow-up are converted to noninterviews. The preliminary percent noninterview is compared to the percent noninterview after the before follow-up matching is complete in the remainder of the tables in this document. The percent noninterview in the following tables is the percentage of noninterviews in occupied P-sample housing units.

Of all interviews at occupied housing units 7.2 percent were proxy interviews and 92.8 percent were interviews with household members.

Table 1.5: Percent Noninterview in the P-sample Before Follow-up by Respondent Type
(Unweighted Data)

Respondent Type	P-sample Percent Noninterview	
	Preliminary Percent Noninterview	Percent Noninterview After Before Follow-up Matching
Household member	0.9	1.7
Proxy	13.8	17.0
Total	1.9	2.8

Table 1.6: Percent Noninterview in the P-sample Before Follow-up by Type of Enumeration Area
(Unweighted Data)

Type of Enumeration Area	P-sample Percent Noninterview		
	Percent of Census Housing Units	Preliminary Percent Noninterview	Percent Noninterview After Before Follow-up Matching
Mail Out/Mail Back	78.7	2.0	2.8
Update/Leave	19.5	1.6	2.7
List/Enumerate	0.3	0.9	1.6
Rural Update/Enumerate	0.7	1.1	3.6
Urban Update/Leave	0.2	1.2	2.7
Urban Update/Enumerate	0.1	2.3	3.5
Adds to Address List	0.4	1.1	2.5
Total		1.9	2.8

Of all interviews, 33.5 percent of the occupied housing unit were completed by telephone, 66.1 percent were completed by personal visit, and 0.3 percent, which is 910 interviews, were completed by a quality assurance replacement interview. The percent noninterview of occupied housing units for the interview mode is in the next table.

Telephone interviews are more likely to have insufficient information because we only have one household to get the information from, there was no opportunity to get better information from a different respondent. Also, there were telephone interviews where we talked to the inmover and they did not have information about the outmover. If the people moved into the address after census day, completed the census questionnaire and mailed it back, we could have called an inmover.

There are several reasons for a high noninterview rate for the quality assurance replacement interviews. These were difficult interviews because they failed the quality assurance check and needed a reinterview. Many of the noninterviews were refusals. Additionally, because the instrument is monitoring both the quality assurance case and the replacement interview, it was difficult to obtain the census day residents in mover cases and many of these were noninterviews. There was also a problem with the instrument in cases where the quality assurance interviewer could not find the address. In these cases, the case failed the quality assurance check but no data was collected for the replacement interview since the address did not exist on the day of the interview. These are considered noninterviews because the cases closed up before any census day information could be obtained.

Table 1.7: Percent Noninterview in the P-sample Before Follow-up by Interview Mode
(Unweighted Data)

Interview Mode	P-sample Percent Noninterview	
	Preliminary Percent Noninterview	Percent Noninterview After Before Follow-up Matching
Telephone	0.9	1.2
Personal Visit	2.2	3.5
Quality Assurance Replacement	36.0	37.4
Total	1.9	2.8

Table 1.8: Percent Noninterview in the P-sample Before Follow-up by Census Regional Office
(Unweighted Data)

Census Regional Office	P-sample Percent Noninterview	
	Preliminary Percent Noninterview	Percent Noninterview After Before Follow-up Matching
Boston	0.2	0.8
New York	2.6	4.4
Philadelphia	2.5	3.6
Detroit	1.6	2.2
Chicago	1.7	2.5
Kansas City	1.5	2.1
Seattle	1.8	2.6
Charlotte	3.2	4.1
Atlanta	2.2	3.3
Dallas	2.1	2.9
Denver	1.4	2.4
Los Angeles	1.9	2.8
Total	1.9	2.8

Table 1.9: Percent Noninterview in the P-sample Before Follow-up by Census Region
(Unweighted Data)

Census Region	P-sample Percent Noninterview	
	Preliminary Percent Noninterview	Percent Noninterview After Before Follow-up Matching
Northeast	1.4	2.5
Midwest	1.6	2.3
South	2.5	3.4
West	1.8	2.6
Total	1.9	2.8

Table 1.10: Percent Noninterview in the P-sample Before Follow-up by State (Unweighted Data)		
State	P-sample Percent Noninterview	
	Preliminary Percent Noninterview	Percent Noninterview After Before Follow-up Matching
Alabama	1.6	2.5
Alaska	3.1	3.4
Arizona	1.7	3.6
Arkansas	0.9	1.7
California	1.8	2.6
Colorado	1.5	1.9
Connecticut	0.1	0.5
Delaware	1.9	2.6
District of Columbia	5.5	7.4
Florida	2.6	3.8
Georgia	1.9	2.7
Hawaii	3.4	4.9
Idaho	1.3	1.9
Illinois	1.5	2.6
Indiana	2.4	3.0
Iowa	1.0	1.6
Kansas	1.5	2.0
Kentucky	3.3	4.0
Louisiana	2.1	3.1
Maine	0.1	0.4

Maryland	3.5	4.7
Massachusetts	0.3	0.9
Michigan	1.6	2.1
Minnesota	1.2	1.7
Mississippi	0.7	1.7
Missouri	2.6	3.4
Montana	0.8	1.8
Nebraska	1.3	1.6
Nevada	2.7	3.7
New Hampshire	0.1	0.6
New Jersey	2.1	3.2
New Mexico	1.1	2.2
New York	2.0	3.6
North Carolina	3.8	5.1
North Dakota	0.9	1.7
Ohio	1.7	2.3
Oklahoma	0.9	1.5
Oregon	2.0	2.6
Pennsylvania	1.5	2.7
Rhode Island	0.5	1.3
South Carolina	2.9	3.6
South Dakota	0.9	1.8
Tennessee	2.8	3.8
Texas	2.3	3.1
Utah	0.9	1.5
Vermont	0.0	1.2

Virginia	2.8	3.7
Washington	1.3	2.1
West Virginia	1.6	2.6
Wisconsin	1.3	1.7
Wyoming	1.8	2.2
Total	1.9	2.8

2. Before Follow-up Code Summary

The first two tables contain the results of before follow-up matching for the P-sample and the E-sample. For details of these codes, see Childers (2000). These before follow-up matching results are from unweighted data from the fifty states and the District of Columbia. These tables do not include the before follow-up matching results in Puerto Rico. The P-sample codes are grouped into

- Matched
- Not matched
- Possible match
- Unresolved match status
- Removed from the P-sample

Matched - The P-sample person is found in the census.

Not Matched - The P-sample person is not found in the census. A follow-up interview is conducted for

- partial household nonmatches
- whole households of conflicting household members (i.e., whole households of P-sample and census nonmatches)[1]
- other whole household nonmatches where the P-sample interview was conducted with a nonhousehold member[2]

Unresolved Match Status - The unresolved before follow-up is insufficient information for matching and follow-up for the P-sample person and possible matches.

Removed from the P-sample - The only category of removed from the P-sample in the before follow-up matching are the P-sample people coded as duplicates. The P-sample duplicates are removed because they are listed more than once.

The E-sample codes are grouped into

- Correctly enumerated

---

[1] These cases have been called the Smith/Jones cases in the past.

[2] No follow-up interview is conducted when there are whole households of P-sample nonmatches from interviews with household members in a housing unit that did not match in the housing unit operation or matched to a housing unit containing no data defined people.

- Erroneously enumerated
- Not matched and needing a follow-up interview
- Possible match
- Unresolved

Correctly enumerated - At this point, the only correctly enumerated people are the ones matching the P-sample.

Erroneously enumerated - The categories during before follow-up are fictitious people, duplicates, insufficient information for matching and follow-up, and geocoding errors.

- The fictitious people are ones where we found notes on the census image identifying the person as not a real person such as a dog or other pet.
- The E-sample people enumerated more than once are coded as duplicates.
- The E-sample people with insufficient information for matching and follow-up are ones who are data defined, but do not contain full name and at least two characteristics.[3]
- Census people in housing units identified as geocoding errors[4] during the housing unit follow-up are coded as erroneously enumerated because of geocoding error.

Unresolved enumeration status - In before follow-up matching, the unresolved category only includes the census housing units needing targeted extended search field work that was not done.

E-sample nonmatches - All E-sample people who do not match to the P-sample are sent for a follow-up interview. The E-sample nonmatches will be coded as correctly or erroneously enumerated or with unresolved enumeration status after follow-up.

E-sample possible matches - E-sample people who were coded as possible matches are followed up to determine whether they are, in fact, matches.

---

[3] This is the same rule that was used in the 1990 PES. There must be enough information about the person to have a chance at locating the person for a follow-up interview before the person is allowed into the matching process. See Childers (2000).

[4] A geocoding error is an error in assigning the housing unit to the correct location.

Table 2.1: National P-Sample Before Follow-up Matching		
P-sample Match Status	Unweighted People	Percent
Matched	573,506	85.7
Not Matched	76,804	11.5
Insufficient Information	7,524	1.1
Possible Match	5,070	0.8
Removed	5,923	0.9
Total	668,827	100.0

Table 2.2: National E-Sample Before Follow-up Matching		
E-sample Enumeration Status	Unweighted People	Percent
Correctly Enumerated	544,995	76.4
Erroneously Enumerated	27,934	3.9
Unresolved	304	0.0
Possible Match	4,751	0.7
Not Matched	134,916	18.9
Total	712,900	100.0

Table 2.3: Erroneous Enumerations in Before Follow-up Matching			
E-sample Erroneous Enumeration Code	Unweighted People	Percent of Erroneous Enumerations	Percent of E-sample
Duplicate	5,511	19.7	0.8
Geocoding Error	9,018	32.3	1.3
Insufficient Information	13,358	47.8	1.9
Fictitious	47	0.2	0.0
Total	27,934	100.0	

Table 2.4 contains the final weighted and imputed data after person follow-up in 1990 for the erroneous enumerations.

Table 2.4: 1990 Erroneous Enumerations Final Weighted Numbers		
E-sample Erroneous Enumeration Code	Percent of Erroneous Enumerations	Percent of E-sample
Duplicate	28.2	1.6
Geocoding Error	6.0	0.3
Insufficient Information	20.8	1.2
Fictitious	2.6	0.2
Other Residence	38.0	2.2
Unresolved	4.5	0.3
Total	100.0	5.8

3. Before Follow-up Results for Location of Matches

Matching is first to the E-sample in the sample block cluster. If the cluster is large, there is matching to the non E-sample people within the cluster. When the search area is expanded to the first ring of blocks surrounding the sample cluster, the P-sample is matched to the census people in the first ring.

Table 3.1: Percent Of Matches in the P-sample by Location of the Census Match Before Follow-up by Type of Enumeration Area
(Unweighted Data)

Type of Enumeration Area	Matches to E-sample	Matches to Non E-sample within Cluster	Matches in the Surrounding Blocks
Mail Out/Mail Back	94.9	0.8	4.3
Update/Leave	96.2	2.1	1.6
List/Enumerate	69.4	30.6	--
Rural Update/Enumerate	98.6	0.4	1.0
Urban Update/Leave	91.4	0.4	8.2
Urban Update/Enumerate	99.9	0.0	0.1
Adds to Address List	97.4	1.9	0.7
Total	95.0	1.2	3.8

Table 3.2: Percent Of Matches in the P-sample by Location of the Census Match Before Follow-up by Tenure
(Unweighted Data)

Tenure	Matches to E-sample	Matches to Non E-sample within Cluster	Matches in the Surrounding Blocks
Owner	96.2	0.9	2.9
Renter	92.4	2.0	5.6
Blank	95.8	1.3	2.8
Total	95.0	1.2	3.8

Table 3.3: Percent Of Matches in the P-sample by Location of the Census Match Before Follow-up by Region
(Unweighted Data)

Region	Matches to E-sample	Matches to Non E-sample within Cluster	Matches in the Surrounding Blocks
Northeast	94.9	1.0	4.1
Midwest	96.7	0.5	2.8
South	94.4	1.3	4.3
West	94.5	1.8	3.7
Total	95.0	1.2	3.8

4. Before Follow-up Percent Not Matched

The P-sample nonmatch rate is calculated by dividing the unweighted number of P-sample nonmatches by the unweighted P-sample total. This P-sample total does not include the people coded as removed. The E-sample nonmatch rate is also the unweighted number of E-sample nonmatches divided by the total unweighted E-sample.

Table 4.1: Percent Not Matched Before Follow-up by Sex
(Unweighted Data)

Sex	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
Male	12.3	19.4	-7.1
Female	10.8	18.5	-7.7
Blank	14.5	18.7	-4.2
Total	11.6	18.9	-7.3

Table 4.2: Percent Not Matched Before Follow-up by Age
(Unweighted Data)

Age	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
Under 18	11.9	18.2	-6.3
18 to 29	16.3	25.6	-9.3
30 to 49	10.9	17.4	-6.5
Over 50	9.5	16.9	-7.4
Blank	14.7	23.0	-8.3
Total	11.6	18.9	-7.3

Table 4.3: Percent Not Matched Before Follow-up by Race
(Unweighted Data)

Race	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
White	9.7	17.3	-7.6
Black	15.8	24.1	-8.3
American Indian	20.9	19.4	-1.5
Asian	12.1	21.0	-8.9
Native Hawaiian and Pacific Islander	16.9	29.6	-12.7
Other Race	14.7	24.3	-9.6
Multiple Race	13.0	20.4	-7.4
Blank	18.4	19.6	-1.2
Total	11.6	18.9	-7.3

Table 4.4: Percent Not Matched Before Follow-up by Hispanic Origin
(Unweighted Data)

Hispanic Origin	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
Hispanic	14.4	22.6	-8.2
Non-Hispanic	11.0	18.3	-7.3
Blank	17.5	19.8	-2.3
Total	11.6	18.9	-7.3

Table 4.5: Percent Not Matched Before Follow-up by Tenure
(Unweighted Data)

Tenure	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
Owner	9.2	14.5	-5.3
Renter	16.1	27.1	-11.0
Blank	16.0	22.4	-6.4
Total	11.6	18.9	-7.3

Table 4.6: Percent Not Matched Before Follow-up by Type of Enumeration Area
(Unweighted Data)

Type of Enumeration Area	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
Mail Out/Mail Back	11.2	19.1	-7.9
Update/Leave	11.4	16.8	-5.4
List/Enumerate	18.3	42.2	-23.9
Rural Update/Enumerate	20.7	18.4	+2.3
Urban Update/Leave	10.9	23.2	-12.3
Urban Update/Enumerate	13.4	13.1	+0.3
Adds to Address List	17.6	21.9	-4.3
Total	11.6	18.9	-7.3

Table 4.7: Percent Not Matched Before Follow-up by Census Region
(Unweighted Data)

Census Region	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
Northeast	12.0	19.1	-7.1
Midwest	8.9	13.2	-4.3
South	12.6	22.4	-9.8
West	12.3	19.1	-6.8
Total	11.6	18.9	-7.3

Table 4.8: Percent Not Matched Before Follow-up by Census Regional Office
(Unweighted Data)

Census Regional Office	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
Boston	11.8	17.4	-5.6
New York	13.6	23.3	-9.7
Philadelphia	12.2	20.6	-8.4
Detroit	8.2	12.8	-4.6
Chicago	10.3	14.3	-4.0
Kansas City	8.4	13.3	-4.9
Seattle	11.3	18.0	-6.7
Charlotte	11.9	22.0	-10.1
Atlanta	12.9	20.8	-7.9
Dallas	13.1	24.3	-11.2
Denver	13.4	17.8	-4.4
Los Angeles	11.4	20.2	-8.8
Total	11.6	18.9	-7.3

Table 4.9: Percent Not Matched Before Follow-up by Percent Mobile Home
(Unweighted Data)

Percent Mobile Home	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
None	11.0	18.7	-7.7
10 Percent or less	11.4	16.7	-5.3
11 to 50 percent	13.8	19.5	-5.7
Greater than 50 percent	17.3	26.5	-9.2
Total	11.6	18.9	-7.3

Table 4.10: Percent Not Matched Before Follow-up by Percent Multi-Unit
(Unweighted Data)

Percent Multi-Unit	P-sample Percent Not Matched	E-sample Percent Not Matched	Difference
None	10.7	17.4	-6.7
10 Percent or less	9.7	16.3	-6.6
11 to 50 percent	11.0	17.9	-6.9
Greater than 50 percent	15.0	24.4	-9.4
Total	11.6	18.9	-7.3

The next set of tables contain percent not matched in the P-sample for variables that are only on the P-sample data.

Table 4.11: Percent Not Matched in the P-sample Before Follow-up by Type of Address (Unweighted Data)	
Type of Address	P-sample Percent Not Matched
Single Unit	9.7
Multi-Unit	16.2
Mobile Home not in a Park	16.7
Mobile Home in a Park	20.2
Single Housing Unit in a Special Place	27.2
Multi-Unit in a Special Place	35.7
Other	45.9
Total	11.6

The next table contains the P-sample percent not matched for the original classification of the P-sample housing unit listed in the Fall of 1999. For example, a P-sample housing unit may have been listed as future construction in the Fall of 1999 and existed as a housing unit during the housing unit follow-up operation. An interview was conducted at the housing unit and 38.6 of the P-sample census day residents were not found in the census within the search area.

Table 4.12: Percent Not Matched in the P-sample Before Follow-up by Type of Unit at Listing (Unweighted Data)	
Type of Unit at Listing	P-sample Percent Not Matched
Housing Unit	11.4
Under Construction	29.7
Future Construction	38.6
Unfit for Habitation	46.3
Boarded Up	29.2
Storage	23.1
Vacant Trailer Site	23.3
Other	70.3
Total	11.6

The P-sample people are classified as nonmovers, outmovers, and people with unresolved residence status. The P-sample is 95.2 percent nonmovers, 3.4 percent outmovers, and 1.4 percent unresolved residence status. The percent not matched by mover status is in the next table.

Table 4.13: Percent Not Matched in the P-sample Before Follow-up by Mover Status (Unweighted Data)	
Mover Status	P-sample Percent Not Matched
Nonmover	10.9
Outmover	26.6
Unresolved Mover Status	23.6
Total	11.6

Table 4.14: Percent Not Matched in the P-sample Before Follow-up by Housing Unit Match Status
(Unweighted Data)

Housing Unit Match Status	P-sample Percent Not Matched
Matched	8.9
Not Matched	39.3
Total	11.6

Table 4.15: Percent Not Matched in the P-sample Before Follow-up by Interview Respondent
(Unweighted Data)

Interview Respondent	P-sample Percent Not Matched
Household Member	10.9
Proxy	24.9
Total	11.6

Table 4.16: Percent Not Matched in the P-sample Before Follow-up by Interview Mode
(Unweighted Data)

Interview Mode	P-sample Percent Not Matched
Telephone	2.1
Personal Visit	16.5
Quality Assurance	16.9
Total	11.6

5. Before Follow-up Results for Census Duplicates

In 1990, 1.6 percent of the census was estimated to be duplicated.

The numbers in this document are unweighted E-sample people coded as duplicate in the 2000 ACE. Duplication is identified in the census in the following situations:

- Between two E-sample people
- Between an E-sample person and a non E-sample person within the sample cluster. The non E-sample people have been subsampled out of the E-sample in large block clusters.
- Between E-sample and non E-sample people in surrounding blocks in TES clusters.

An E-sample person duplicated with another E-sample person is coded as a duplicate and is a full erroneous enumeration. E-sample people are also compared to the non E-sample census people and the probability of erroneous enumeration is estimated. Duplication is also identified between the E-sample people and census enumerations in surrounding blocks in TES clusters. When an E-sample person lives in a household that should have been enumerated in a surrounding block, a search in the surrounding blocks is conducted to identify people duplicated because of geocoding errors. A duplicate code assigned to an E-sample person indicates a full erroneous enumeration. The next several tables contain the duplicate rates in the E-sample using unweighted data.

Table 5.1: Percent Duplication in the E-sample Before Follow-up by Sex
(Unweighted Data)

Sex	E-sample Percent Duplicate
Male	0.8
Female	0.8
Blank	1.6
Total	0.8

Table 5.2: Percent Duplication in the E-sample Before Follow-up by Age (Unweighted Data)	
Age	E-sample Percent Duplicate
Under 17	0.6
18 to 29	0.8
30 to 49	0.7
Over 50	0.7
Blank	3.4
Total	0.8

Table 5.3: Percent Duplication in the E-sample Before Follow-up by Race (Unweighted Data)	
Race	E-sample Percent Duplicate
White	0.7
Black	1.2
American Indian	0.7
Asian	0.9
Native Hawaiian and Pacific Islander	0.9
Other Race	1.4
Multiple Race	0.6
Blank	1.0
Total	0.8

Table 5.4: Percent Duplication in the E-sample Before Follow-up by Hispanic Origin
(Unweighted Data)

Hispanic Origin	E-sample Duplicate
Hispanic	1.1
Non-Hispanic	0.7
Blank	1.3
Total	0.8

Table 5.5: Percent Duplication in the E-sample Before Follow-up by Tenure
(Unweighted Data)

Tenure	E-sample Percent Duplicate
Owner	0.6
Renter	1.1
Blank	1.9
Total	0.8

Table 5.6: Percent Duplication in the E-sample Before Follow-up by Type of Enumeration Area
(Unweighted Data)

Type of Enumeration Area	E-sample Percent Duplicate
Mail Out/Mail Back	0.8
Update/Leave	1.1
List/Enumerate	0.3
Rural Update/Enumerate	0.5
Urban Update/Leave	0.3
Urban Update/Enumerate	0.0
Adds to Address List	0.6
Total	0.8

Table 5.7: Percent Duplication in the E-sample Before Follow-up by Census Region
(Unweighted Data)

Census Region	E-sample Percent Duplicated
Northeast	1.3
Midwest	0.6
South	0.7
West	0.6
Total	0.8

Table 5.8: Percent Duplication in the E-sample Before Follow-up by Census Regional Office (Unweighted Data)	
Census Regional Office	E-sample Percent Duplicated
Boston	1.1
New York	2.2
Philadelphia	0.6
Detroit	0.4
Chicago	0.8
Kansas City	0.6
Seattle	0.8
Charlotte	0.7
Atlanta	0.7
Dallas	0.8
Denver	0.6
Los Angeles	0.6
Total	0.8

Table 5.9 contains the percent person duplication in the E-sample before follow-up by state and the percent housing unit duplication after housing unit matching. The housing unit matching in the Spring of 2000 was a match between the housing units listed in the ACE within cluster to the January DMAF. The percent duplication in the census within cluster is unweighted data after the housing unit matching was completed. The possible duplicates are follow-up to identify real duplicates. The percent duplication for people is unweighted data.

Table 5.9: Comparison of Percent Person and Housing Unit Duplication by State
(Unweighted Data)

State	E-sample Person Percent Duplicate	Census Housing Unit Percent Duplicate
Alabama	1.3	2.4
Alaska	0.6	1.3
Arizona	0.8	0.8
Arkansas	0.7	0.7
California	0.6	0.6
Colorado	0.2	3.7
Connecticut	1.0	2.9
Delaware	1.3	2.0
District of Columbia	1.1	0.2
Florida	0.5	1.4
Georgia	0.7	2.2
Hawaii	1.0	2.8
Idaho	1.3	2.3
Illinois	1.1	1.0
Indiana	0.3	0.4
Iowa	0.6	2.0
Kansas	0.5	0.5
Kentucky	0.8	1.9
Louisiana	1.2	2.8
Maine	1.2	1.6
Maryland	0.5	1.5

Massachusetts	1.1	2.3
Michigan	0.4	0.7
Minnesota	0.3	0.7
Mississippi	0.8	1.3
Missouri	0.9	0.6
Montana	0.4	0.9
Nebraska	0.5	1.4
Nevada	0.4	0.7
New Hampshire	0.5	0.8
New Jersey	1.2	1.0
New Mexico	0.7	1.0
New York	1.9	2.4
North Carolina	0.7	1.7
North Dakota	0.4	0.8
Ohio	0.4	0.7
Oklahoma	0.6	0.5
Oregon	0.5	1.4
Pennsylvania	0.6	1.4
Rhode Island	2.1	1.0
South Carolina	1.0	4.9
South Dakota	0.7	0.3
Tennessee	0.5	1.1
Texas	0.7	0.9
Utah	0.6	1.6
Vermont	2.5	8.3
Virginia	0.4	0.6

Washington	0.8	2.0
West Virginia	0.6	1.2
Wisconsin	0.6	1.5
Wyoming	0.8	0.3
Total	0.8	1.4

Table 5.10: Percent Duplication Before Follow-up by Percent Mobile Home
(Unweighted Data)

Percent Mobile Home	E-sample Percent Duplicate
None	0.7
10 Percent or less	0.8
11 to 50 percent	1.0
Greater than 50 percent	1.2
Total	0.8

Table 5.11: Percent Duplication Before Follow-up by Percent Multi-Unit
(Unweighted Data)

Percent Multi-Unit	E-sample Percent Duplicate
None	0.5
10 Percent or less	0.6
11 to 50 percent	1.0
Greater than 50 percent	1.4
Total	0.8

6. Before Follow-up Results for Census Geocoding Errors

A E-sample person is coded as erroneously enumerated because the housing unit is a geocoding error, i.e., the housing unit's census geography was incorrectly coded. The geocoding error code was applied to the housing unit after the housing unit follow-up in the Spring of 2000. Additional gecoding errors will be discovered during the person follow-up for clusters that did not get matched during housing unit matching, which are the relisted and list/enumerate clusters. Housing units added to the census will also have geography work during the person follow-up interview.

Table 6.1: Percent Geocoding Error in the E-sample Before Follow-up by Sex
(Unweighted Data)

Sex	E-sample Percent Geocoding Error
Male	1.3
Female	1.3
Blank	1.1
Total	1.3

Table 6.2: Percent Geocoding Error in the E-sample Before Follow-up by Age (Unweighted Data)	
Age	E-sample Percent Geocoding Error
Under 17	1.3
18 to 29	1.3
30 to 49	1.3
Over 50	1.2
Blank	0.9
Total	1.3

Table 6.3: Percent Geocoding Error in the E-sample Before Follow-up by Race (Unweighted Data)	
Race	E-sample Percent Geocoding Error
White	1.3
Black	1.0
American Indian	2.5
Asian	1.2
Native Hawaiian and Pacific Islander	0.6
Other Race	1.1
Multiple Race	1.2
Blank	1.2
Total	1.3

Table 6.4: Percent Geocoding Error in the E-sample Before Follow-up by Hispanic Origin (Unweighted Data)	
Hispanic Origin	E-sample Geocoding Error
Hispanic	1.1
Non-Hispanic	1.3
Blank	1.1
Total	1.3

Table 6.5: Percent Geocoding Error in the E-sample Before Follow-up by Tenure (Unweighted Data)	
Tenure	E-sample Percent Geocoding Error
Owner	1.2
Renter	1.4
Blank	1.1
Total	1.3

Table 6.6: Percent Geocoding Error in the E-sample Before Follow-up by Type of Enumeration Area
(Unweighted Data)

Type of Enumeration Area	E-sample Percent Geocoding Error
Mail Out/Mail Back	1.4
Update/Leave	0.5
List/Enumerate	--
Rural Update/Enumerate	2.4
Urban Update/Leave	0.4
Urban Update/Enumerate	1.4
Adds to Address List	0.5
Total	1.3

Table 6.7: Percent Geocoding Error Before Follow-up by Percent Mobile Home
(Unweighted Data)

Percent Mobile Home	E-sample Percent Geocoding Error
None	1.3
10 Percent or less	1.0
11 to 50 percent	1.1
Greater than 50 percent	2.0
Total	1.3

Table 6.8: Percent Geocoding Error Before Follow-up by Percent Multi-Unit (Unweighted Data)	
Percent Multi-Unit	E-sample Percent Geocoding Error
None	1.5
10 Percent or less	0.9
11 to 50 percent	1.1
Greater than 50 percent	0.8
Total	1.3

Table 6.9: Percent Geocoding Error in the E-sample Before Follow-up by Census Region (Unweighted Data)	
Census Region	E-sample Percent Geocoding Error
Northeast	1.6
Midwest	1.1
South	1.2
West	1.3
Total	1.3

Table 6.10: Percent Geocoding Error in the E-sample Before Follow-up by Census Regional Office
(Unweighted Data)

Census Regional Office	E-sample Percent Geocoding Error
Boston	2.0
New York	1.2
Philadelphia	1.5
Detroit	1.0
Chicago	1.1
Kansas City	0.9
Seattle	1.5
Charlotte	1.3
Atlanta	1.4
Dallas	0.9
Denver	1.3
Los Angeles	1.1
Total	1.3

Table 6.11: Percent Geocoding Error in the E-sample Before Follow-up by State (Unweighted Data)	
State	E-sample Percent Geocoding Error
Alabama	0.9
Alaska	0.2
Arizona	1.4
Arkansas	0.5
California	1.2
Colorado	0.6
Connecticut	2.4
Delaware	3.9
District of Columbia	0.0
Florida	1.4
Georgia	1.7
Hawaii	1.6
Idaho	2.0
Illinois	1.1
Indiana	1.4
Iowa	1.8
Kansas	0.5
Kentucky	2.8
Louisiana	1.2
Maine	0.5
Maryland	1.2
Massachusetts	1.9

Michigan	0.8
Minnesota	0.2
Mississippi	1.1
Missouri	1.6
Montana	0.6
Nebraska	0.0
Nevada	0.3
New Hampshire	1.1
New Jersey	1.2
New Mexico	4.2
New York	1.8
North Carolina	1.2
North Dakota	4.4
Ohio	1.2
Oklahoma	0.9
Oregon	1.6
Pennsylvania	1.2
Rhode Island	1.6
South Carolina	1.1
South Dakota	0.3
Tennessee	1.1
Texas	0.8
Utah	0.1
Vermont	1.9
Virginia	1.0
Washington	1.1

West Virginia	1.4
Wisconsin	0.8
Wyoming	0.6
Total	1.3

Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

11/22 ESCAP Presentation

- **Universe=HCUF Prime with 2,374,271 potential deletes excluded**
- **Nationwide, there were 114.9 mill. housing units**
- **This is 0.4 percent less than expected ((HCUF Prime - Hu. Est. 2000)/Hu. Est. April 2000)**
- **Mailout/mailback counties have a shortage of 1.0 percent**
- **Update/leave counties have an overage of 2.7 percent**

Table 1. Difference Between the HCUF PRIME Housing Unit Count and the Housing Unit Estimate for the Nation and for Counties by Type of Enumeration Area

Diff. HCUF -EST.	National	≥ 95% TEA=1	TEA=2 only	TEA=1+2 only	TEA=Mix (excl. 3+4)	TEA=Mix (incl. 3+4)
HCUF Prime	114,900,189	63,095,980	5,319,865	28,579,812	11,940,849	5,963,683
Difference	-495,680	-655,684	138,407	23,948	-3,770	1,419
% Difference	-0.4	-1.0	2.7	0.1	-0.0	0.0
Number of Counties	3,141	390	819	1,301	415	216
Shortage						
HCUF Prime	71,710,976	42,865,252	1,911,899	16,764,094	5,495,783	4,673,948
Difference	1,750,726	985,534	81,567	414,803	191,531	77,291
% Difference	2.4	2.3	4.1	2.4	3.6	1.6
Number of Counties	1,634	266	360	688	210	110
Overage						
HCUF Prime	43,189,213	20,230,728	3,407,966	11,815,718	6,445,066	1,289,735
Difference	1,255,046	329,850	219,974	438,751	187,761	78,710
% Difference	3.0	1.7	6.9	3.9	3.5	6.5
Number of Counties	1,507	124	459	613	205	106

ESCAP MEETING NO. 20 - 11/22/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #20

November 22, 2000

Prepared by: Nick Birnbaum.

The twentieth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on November 22, 2000 at 10:30.

The agenda for the meeting was to discuss how the Census Bureau will release the results of demographic analysis, and to examine results from A.C.E. Before Follow-up Person Matching and preliminary demographic analysis benchmarks of housing units.

Committee Attendees:

Paula Schneider
Nancy Potok
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan
John Long
Susan Miskura
Ruth Ann Killion

Other Attendees:

Kenneth Prewitt
Donna Kostanich
Dan Childers
Debbie Fenstermaker
Kathleen Styles
Maria Urrutia
Signe Wetrogan
Alan Tupek
Jason Devine
Marvin Raines
Tommy Wright
David Whitford
Nick Birnbaum
Carolee Bush
Annette Quinlan
Raj Singh
Kirsten West

I. Availability of Demographic Analysis (DA) Estimates

John Long spoke briefly about the availability of demographic analysis estimates with April 1, 2000, as the reference date. Census level (not adjusted for 1990 undercount) population estimates for April 1, 2000, are currently available to the public (on the Census Bureau's web site). These data are at the national level by age, sex, race, and Hispanic origin.

Additionally, users, if so inclined, could produce national level adjusted population estimates for April 1, 2000, by applying the 1990 PES or the 1990 demographic analysis undercount rates to the census level population estimates mentioned above. These undercount rates are also available on the Census Bureau's web site.

The official demographic analysis population data for April 1, 2000, will be released at the time the Committee delivers its report to the Director, as part of the documentation underlying its analyses and recommendation regarding the use of the adjusted data for redistricting purposes.

II. Demographic Benchmark Analysis of Housing Unit Estimates

Kirsten West and Jason Devine presented preliminary demographic benchmark analysis results for housing units. The demographic benchmark is the housing unit estimate for July 1999 projected to April 1, 2000. The analysis is done for the nation and for groups of counties by type of enumeration area (TEA). For the nation and mailout/mailback counties, the preliminary census file housing unit count was slightly lower than the benchmark. For counties that are solely update/leave, the preliminary census file housing unit count mildly exceeded the benchmark, but it was noted that there is greater uncertainty associated with the DA housing unit estimates for these areas.

III. Results from A.C.E. Before Follow-up (BFU) Person Matching

Dan Childers began his presentation on the results from the BFU Person Matching by reviewing data on P-sample noninterview rates (unweighted data). He also discussed data showing percentages of person and housing unit duplication in the E-sample (unweighted data). Dan will continue his presentation on the BFU Person Matching results at the next meeting.

IV. Next Meeting

The agenda for the next meeting, scheduled for November 30, 2000, is to continue discussions of the results from BFU Person Matching.

ESCAP MEETING NO. 21 - 11/30/00

AGENDA

Kathleen P Porter
11/28/2000 02:01 PM

To: Angela Frazier/DMD/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Danny R Childers/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Kathleen P Porter/DMD/HQ/BOC@BOC
cc:
Subject: Re: Agenda for Nov. 30 ESCAP Meeting

PLEASE REMEMBER TO BRING THE FOLDER THAT YOU RECEIVED AT THE LAST MEETING.

THANKS.

Kathleen P Porter
11/28/2000 10:49 AM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A

Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC, Mary E Williams/DIR/HQ/BOC, Danny R Childers/DSSD/HQ/BOC

cc:

Subject: Agenda for Nov. 30 ESCAP Meeting

* * * * * * * * * * PLEASE NOTE THE DATE AND TIME OF THIS MEETING * * * * * * * * * * * *

The agenda for the November 30 ESCAP Meeting scheduled from 2:30-4:00 in Rm. 2412/3 is as follows:

A.C.E. Before Follow-Up Match Results (continued) - Dan Childers

ESCAP MEETING NO. 21 - 11/30/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 21

November 30, 2000

Prepared by: Maria Urrutia and Annette Quinlan

The twenty first meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on November 30, 2000 at 2:30. The agenda for the meeting was to continue the discussion of the A.C.E. Before Followup Match results.

Committee Attendees:

William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long

Other Attendees:

Kenneth Prewitt	Debbie Fenstermaker
Marvin Raines	Roxie Jones
Donna Kostanich	Kathleen Styles
Raj Singh	Maria Urrutia
David Whitford	Annette Quinlan
Danny Childers	

I. Before Followup Match Results

Danny Childers continued the Before Followup (BFU) matching results discussion from the previous ESCAP meeting (November 22). All results discussed are the preliminary, unweighted results and don't include the Targeted Extended Search results. These data were discussed in the context of how they compared to 1990 results, and as indications of gross trends that might result in the final DSE data.

The preliminary BFU matching results for the P and E samples by different geographic areas, including Type of Enumeration Area (TEA) and region, were presented to the ESCAP. These data were discussed, and it was noted that, for the P-sample, they compared reasonably with analogous 1990 results.

The preliminary E-sample BFU erroneous enumeration results were also examined, but only final 1990 results were available, so a full comparison could not be made. It was possible to compare the geocoding error component with 1990 results. This comparison indicated that the level of geocoding error for 2000 was likely to be larger than in 1990. The effect of greater geocoding error was discussed, and it was noted that geocoding errors, if clustered, would result in increased variances. Therefore, we will carefully consider the estimates of variance that will be produced for the DSEs.

Another finding for the E-sample was the low percentage of people counted at the wrong location. Howard Hogan and his staff are examining potential causes for this finding, and will report back at a subsequent meeting.

The ESCAP also considered the BFU E- and P-sample results in terms of how they might predict final coverage errors. The data did not allow for any quantitative assessments, but did indicate that the minority post-strata were likely to have higher undercount rates than the non-minority post-strata. However, this comparison was very rough and could not be put into quantitative measurements.

II. Next Meeting

The next meeting scheduled for Wednesday December 6, 2000 will discuss Service Based Enumeration quality indicators.

ESCAP MEETING NO. 22 - 12/06/00

AGENDA

Preston J Waite
Sent by: Sue A Kent
12/05/2000 03:36 PM

To: Angela Frazier/DMD/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Danny R Childers/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Kathleen P Porter/DMD/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC
cc: Kathleen P Porter/DMD/HQ/BOC@BOC

Subject: 12-6 ESCAP Meeting Agenda (Noon - 1:30 p.m.)

The agenda for tomorrow's ESCAP Meeting (12:00 - 1:30 p.m.), in Room 2412/3 follows:

Preview of SBE Quality Indicators - Rick Griffin

Attached is an overview memorandum on SBE for your review.

ESCAP MEETING NO. 22 - 12/06/00

HANDOUTS

UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

December 5, 2000

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES #Q-36

MEMORANDUM FOR Howard Hogan
Chief, Decennial Statistical Studies Division

From: Donna Kostanich *(signed 12/5/2000)*
Assistant Division Chief, Sampling and Estimation
Decennial Statistical Studies Division

Prepared by: Felipe Kohn
Estimation Staff
Decennial Statistical Studies Division

Subject: Census 2000 Service Based Enumeration: Overview of Multiplicity Estimation

I. INTRODUCTION

A key component of Census 2000 is the enumeration of persons with no usual residence. These persons had the opportunity to be enumerated during Service Based Enumeration (SBE) at shelters, soup kitchens, mobile food vans and Targeted Non-sheltered Outdoor Locations (TNSOLS). In addition, persons with no usual residence had the opportunity to be enumerated by completing Be Counted Forms (BCF). For the uncorrected Census count, the Census Bureau will include the persons actually enumerated (after unduplication) at shelters, soup kitchens, mobile food vans, TNSOLS and on BCF forms indicating no usual residence as the official count.

For the corrected Census count the Bureau plans to use the multiplicity estimator (based on the service usage questions asked in the questionnaires) to estimate the number of persons without usual residence who use shelters, soup kitchens and mobile food vans (SBE facilities). In addition, persons enumerated in TNSOLS and persons without usual residence that filed BCFs will augment the estimate. Since each of the SBE facilities is enumerated on only one day, the multiplicity estimator uses frequency of facility usage responses to estimate the number of persons using services but not on the day of enumeration. The multiplicity estimator is a statistical technique based on the service usage questions asked in the Individual Census Report (ICR) used in shelters, and the Individual Census Questionnaire (ICQ) used in soup kitchens and other facilities.

This memorandum provides an overview of the SBE estimation procedures. For more detail see the memorandum in Chapter Q of this series, Subject: Census 2000 Service-Based Enumeration: Computer Specifications for Multiplicity Estimation. Multiplicity Estimation is done independently for each county.

II ENUMERATION

Census Bureau enumerators visited shelters to collect information on the 27th of March and the following day went to collect information in soup kitchens and mobile food vans.

For the Census 2000 SBE enumeration, the Bureau used four forms: two for the population found in shelters and two for the population found in soup kitchens and mobile food vans. Questions were asked about shelter usage and soup kitchen or mobile food van usage. The responses to the shelter usage question and the soup kitchen usage question whenever the respondent answers "No" to the shelter usage question will be used for production estimation for the SBE population in the 2000 Census.

Shelter Usage

For the shelter population, the Bureau used the ICR short form as well as long form versions. In these questionnaires Question 2 d asks the following shelter question: "Including tonight, how many nights during the past 7 nights did you stay in a SHELTER?". The ICQ short and long versions given to persons in soup kitchens and using mobile food vans also ask clients about their shelter usage. Question 10 in the short version and question 9 in the long version ask "Including last night, during the past seven nights did you stay in a SHELTER?"(see summary table below). In the ICR questionnaire the respondent is asked about his/her usual place of residence. This question is on the ICR because this form is used for other data collection operations. However, for SBE we assume that all persons staying in a shelter should have the shelter's address as their usual place of residence regardless of their response to the usual place of residence question.

Soup Kitchen or Mobile Food Van Usage

The soup kitchen or mobile food van usage question is limited to the ICQ questionnaire short and long versions. Question 9 in the short and Question 8 in the long ask the following question: "Including today, how many days during the past seven days did you receive a meal from a SOUP KITCHEN or MOBILE FOOD VAN?"(see summary table below).

Each person is also asked if he/she has a usual residence. Any person who responds that he/she has a usual residence is removed from the file.

The ICR questionnaire is used for the TNSOL enumeration, but persons enumerated there are not asked the shelter usage question or the usual place of residence question.
In addition, persons without usual residence could fill and return a BCF. The BCF does not include any service usage question.

Table 1: Summary of Service Based Enumeration

Enumerated Population	**Date**	**Questionnaire Used**	**Usage Questions**
Shelter	March 27th	ICR Short & Long	"Including tonight, how many nights during the past seven nights did you stay in a shelter?"
Soup Kitchen or Mobile Food Van	March 28th	ICQ Short & Long	"Including today how many days during the past seven days did you receive a meal from a SOUP KITCHEN or MOBILE FOOD VAN?" "Including last night, during the past seven nights did you stay in a SHELTER?"
Targeted Non-Sheltered Outdoors Locations (TNSOLS)	March 31st	ICR Short	Not Asked

III. RULES OF UNDUPLICATION

The persons without usual residence that use service facilities are very transient by definition. Since the enumeration of shelters was done the 27th of March, and the enumeration of soup kitchens was done the following day, it is expected that at least some of the persons were enumerated more than once. In the paragraph below, we describe the rule we used for ensuring one (and only one) record for each person.

If a person is enumerated in a shelter and in a soup kitchen or food van, then when unduplicating the shelter questionnaire should be kept. If a person is enumerated in a service

facility and in a TNSOL, when unduplicating the form collected in the service facility should be kept. Finally, if a person is enumerated in SBE and on a BCF with no usual residence, the SBE questionnaire will be kept (see summary table below).

Table 2: Summary of Unduplications

Duplicate Enumerations		Questionnaire Kept
Shelter	Soup Kitchen	Shelter
Shelter	TNSOL	Shelter
Shelter	BCF	Shelter
Soup Kitchen	TNSOL	Soup Kitchen
Soup Kitchen	BCF	Soup Kitchen
TNSOL	BCF	TNSOL

IV. IMPUTATION

If a questionnaire has some missing demographic data the Population Division imputes that missing data item in the same way as in housing questionnaires. Imputation of missing response to the usage questions on SBE questionnaires is done as part of SBE estimation prior to multiplicity estimation. For each state, the mean usage of respondents will be imputed on the input file (Hundred percent Census Edited File, HCEF) for non-respondents within age/sex/type of facility (shelter or soup kitchen/mobile food van) groups prior to multiplicity estimation. Imputation cells with less than 10 respondents will be collapsed. For soup kitchen and mobile food van questionnaires, no imputation of the soup kitchen/mobile food van question is necessary if the response to the shelter usage question is "YES".

V. MULTIPLICITY ESTIMATION

A. If the usage question used in estimation had to be imputed for all persons in a shelter, soup kitchen, or mobile food van, then all persons at the SBE site are given a weight of one (i.e., no multiplicity estimation for the site).

B. Assigning Weights

1. Shelter Persons

After the unduplication of enumerated persons in shelters, the initial weight for all shelter person records will be calculated using the formula below:

$W_j = \frac{7}{A_j}$ where A_j is the response or imputed value for the ICR shelter usage question for the j-th person enumerated in a shelter .

2. Soup Kitchens and Mobile Food Vans persons

For all ICQ records (after unduplication) if their response to the shelter usage question is "YES" they are given a weight of zero for the multiplicity estimator. These persons are accounted for in the multiplicity factor applied to shelter respondents.

If the response to the shelter question is "NO" or there is "no response", then after the unduplication of enumerated persons in soup kitchens and mobile food vans, the initial weight for all soup kitchen and mobile food van records will be calculated using the formula below:

$W_j = \frac{7}{B_j}$ *where,* B_j is the response or imputed value for the ICQ soup kitchen usage response for the j-th person enumerated in a soup kitchen or mobile food van.

3. TNSOLS and BCF with no usual residence have an effective weight of one and are not subject to the multiplicity estimation process.

Thus, the multiplicity estimator of persons without usual residence that use services has the formula shown below:

$$\hat{X} = \sum_{j=1}^{n} \frac{7}{A_j} + \sum_{j=1}^{m} \frac{7}{B_j} + X_{TNSOL} + X_{BCF}$$

Where, after unduplication, n is the number of persons enumerated at shelters and m is the number of persons enumerated at soup kitchens or mobile food vans who did not answer "YES" to the shelter usage question.

VI. RECORD CREATION

The output file from multiplicity estimation will be the HEDF with SBE persons WITHOUT WEIGHTS. Thus, the multiplicity weights described in Section V. are reduced by one and control rounded to integers (the sum of the integer weights will be within 1 of the sum of the unrounded weights) to produce rounded replication weights. Any person who matched to a person enumerated at a TNSOL or BCF with no usual residence is given a replication weight of zero (i.e., they represent only themselves and no additional people) to account for other TNSOL or BCF persons who would have matched if SBE was done on a different day(s) of the week. For the final HEDF all SBE enumerated persons after unduplication remain on the file and are duplicated a number of times equal to these rounded replication weights.

In addition we do not want to remove persons from the file who have a weight of zero in the multiplicity estimator since they have taken the time to supply responses to demographic data. These are the persons enumerated at soup kitchens and mobile food vans who responded "YES" to the shelter usage question. If the multiplicity estimator turns out to be less than the uncorrected count this is indicative of considerable response bias. Thus, we use the following rule for each county.

Count 1: The number of persons added to the uncorrected count due to multiplicity estimation.

Count 2: The number of persons enumerated at soup kitchens and mobile food vans who responded "YES" to the shelter usage question (these are the persons given a weight of zero for the multiplicity estimate).

If, Count 2 is greater than Count 1 then multiplicity estimation will not be done for the county (since the multiplicity estimation process would produce a count lower than the uncorrected count) . The final count for SBE for the corrected Census count file will be the same as for the uncorrected Census count (i.e., the number of persons after unduplication enumerated at shelters, soup kitchens, mobile food vans, TNSOLS, or BCFs without usual residence).

If, Count 1 is greater than or equal to Count 2, a sample of size s = Count 2 of the persons with rounded replication weights not equal to zero will be selected allowing persons to be selected more than once with probability proportional to their replication weights. Each selected person will have their rounded replication weight reduced by 1 for each time they are selected in the sample (i.e., a person selected twice will have their rounded replication weight reduced by 2). Here we are reducing the number of duplications so that we can leave persons with a weight of zero on the file. For each person with a weight of zero, we reduce the number of duplications by one.

VII. REFERENCES

Sirken, M.G. (1970), "Households Surveys with Multiplicity", Journal of the American Statistical Association, Volume 63, 257-266.

Sirken, M.G. (1972), "Stratified Sample Surveys with Multiplicity", Journal of the American Statistical Association, Volume 67, 224-227.

Sirken, M.G. (1972), "Variance Components of Multiplicity Estimators", Biometrics, Volume 28, 869-873.

Sirken, M.G. and Leoy, P.S. (1974), "Multiplicity Estimation of Proportions based on Ratios of Random Variables"

Thompson, S.K., (1992), "Sampling", John Wiley and Sons, Inc.

cc: Census 2000 Procedures and Operations Memorandum Distribution List

COMPARISON OF 2000 SBE AND 1990 S-NIGHT
Number of People Tabulated at Each Location

DEC 0 6 2000

FIPS	State	Total SBE POP 2000	Total SBE W/out Soup Kit & Vans	Total S-Night POP	Diff Bet Total SBE 2000 & 1990	Diff SBE (w/out Soup Kit)/ Vans) & 1990
	TOT	280527	206494	240140	16 82	-14 91

INFORMATION DELETED

Outdoor Locations			Other Services		
2000 POP	1990 POP	Diff Bet 2000 & 1990	Soup Kitchens	Mobile Vans	Total SK & Vans
23080	49734	-53 59	70604	3429	74033
165	364	-54 67			
25	79	-68 35			
1079	1897	-43 12			
74	62	19 35			
7635	18081	-57 77			
221	393	-43 77			
49	221	-77 83			
8	19	-57 89			
162	131	23 66			
2228	3189	-30 13			
224	450	-50 22			
527	1071	-50 79			
8	19	-57 89			
258	1755	-85 30			
678	268	152 99			
61	148	-58 78			
4	156	-97 47			
217	116	83 90			
81	184	-55 98			
0	7	-100 00			
222	523	-57.55			
94	674	-86.05			
2588	262	887 79			
30	138	-78 26			
21	83	-74 70			
42	215	-80 47			
16	17	-5.88			
9	20	-55.00			
609	436	39 68			
12	8	50 00			
604	1639	-63 15			
215	164	31.10			
1239	10732	-88 46			
184	259	-28 96			
0	30	-100.00			
213	188	13 30			
21	340	-93 82			
248	564	-56.03			
239	1312	-81 78			
5	44	-88 64			
102	102	0 00			
11	71	-84 51			
217	357	-39 22			
1202	1442	-16 64			
244	276	-11 59			
30	16	87 50			
171	319	-46 39			
747	772	-3.24			
27	33	-18 18			
13	71	-81 69			
1	13	-92 31			
345	450	-23 33			

ESCAP MEETING NO. 22 - 12/06/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 22
December 6, 2000

Prepared by: Nick Birnbaum.

The twenty-second meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on December 6, 2000 at 12:00.

The agenda for the meeting was to discuss preliminary data from the Service Based Enumeration.

Committee Attendees:

Nancy Potok
Paula Schneider
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
William Barron
Nancy Gordon
Susan Miskura
John Long
Ruth Ann Killion

Other Attendees:

Marvin Raines
Donna Kostanich
Raj Singh
Debbie Fenstermaker
Roxie Jones
Nick Birnbaum
Carolee Bush
Kathleen Styles
Maria Urrutia
Annette Quinlan
Rick Griffin
Annetta Clark Smith
Denise Smith
Felipe Kohn

I. Discussion of Service Based Enumeration Procedures and Preliminary Data

DSSD staff provided the Committee with some background information regarding the Service Based Enumeration, particularly the use of the multiplicity estimator. The source of this information is DSSD Census 2000 Procedures and Operations Memorandum Series #Q-36, "Census 2000 Service Based Enumeration: Overview of Multiplicity Estimation." This document provides an overview of the SBE estimation procedures and is attached.

DSSD has examined both preliminary data from the actual counts from the SBE and related operations including Targeted Non-sheltered Outdoor Locations and Be Counted Forms indicating no usual residence, and the estimates using the multiplicity estimator for the SBE facilities. The actual counts (unweighted) will be included in the apportionment tabulations.

DSSD staff explained how the estimator is computed and walked through the shelter and soup kitchen usage questions. Respondents' answers are weighted to reflect the fact that the enumeration of SBE facilities only takes place on one day and one has to account, through estimation, for those service users who didn't utilize the services on that particular day. Preliminary SBE data were then presented to the Committee.

DSSD staff expressed concerns about the response patterns to the usage questions. Consequently, the Committee requested more detailed information regarding the usage patterns.

The Committee also reviewed unweighted Census 2000 SBE data in comparison to similar data from the 1990 S-Night operation.

II. Next Meeting

The next meeting, to be held on December 13, will examine more detailed demographic analysis estimates.

Attachment

ESCAP MEETING NO. 23 - 12/13/00

AGENDA

Kathleen P Porter
12/12/2000 02:02 PM

To: Margaret A Applekamp/DIR/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Geneva A Burns/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Angela Frazier/DMD/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Jeannette D Greene/DIR/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, Susan Miskura/DMD/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Kenneth Prewitt/DIR/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Paula J Schneider/DIR/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Ellen Lee/DIR/HQ/BOC@BOC, Annette M Quinlan/DMD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC

cc: J Gregory Robinson/POP/HQ/BOC@BOC

Subject: Agenda for December 13 ESCAP meeting

The agenda for the December 13 ESCAP Meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

Demographic Analysis - Gregg Robinson
Weight Trimming - Howard Hogan

ESCAP MEETING NO. 23 - 12/13/00

HANDOUTS

Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

Presentation for ESCAP on Demographic Analysis (DA)
December 13, 2000

Overall Coverage Levels

* The Census 2000 count of 281.4 million is 1.5 million higher than the preliminary DA estimate of 279.9 million (Table A).

* Relative to DA, this difference implies a net overcount of 0.5 percent.

* This net coverage is dramatically different from that in the 1990 census, where the net undercount was 4.7 million or 1.8 percent (Table 1).

* The DA estimate may be revised upward based on subsequent research (especially the components of undocumented immigration and nonimmigrants)

* In a similar result, the initial 1980 DA estimate implied a net overcount of 0.4 percent, which was later revised upward (in particular, the component of undocumented immigration was added). (Table B).

Coverage Differentials

* The estimated net undercount of males (0.0 percent) is greater than for females (-1.1), but the differential is reduced from 1990 (Table 1).

* Compared to 1990, estimated net undercount is lower in 2000 for all age-sex groups (Table 2).

* Two estimates of net undercount are provided for race groups: Model 1 uses census tabulations for Blacks that include persons who marked the Black circle and no other response circle to the race question; Model 2 uses census tabulations for Blacks that include persons who marked the Black circle and other response circles.

* The alternative DA estimates for Blacks differ (Model 1 estimate of 7.7 percent net undercount, Model 2 estimate of 3.0 percent). The variation is widest for ages 0-17 (7.9 versus -0.1 for Black males). (Table 3 and 4).

* All estimates for 2000 show the highest net undercount for adult Black men.

* Sex ratio for 2000 confirm the differential undercount of Black men relative to Black women; the differential is slightly lower than in 1990 (Table 5 and 6).

PRELIMINARY DEMOGRAPHIC ANALYSIS ESTIMATES for 4-1-2000:
Comparison to HCEF Results (Tabulations based on 'DPrime' File)

AGE	Demographic Analysis	2000 Census	Difference Amount	Difference Percent
TOTAL POPULATION	279,935,096	281,421,906	(1,486,810)	-0.5
Male	138,064,987	138,053,563	11,424	0.0
Female	141,870,109	143,368,343	(1,498,234)	-1.1

Note: The DA estimates for ages 65+ in 2000 are based on DA estimates for 55+ in 1990 carried forward with components of change. Revised DA estimates for 65+ in 2000 will be based on 2000 Medicare data. The estimates for non-immigrants used to develop the total DA estimate will be revised using current data on foreign students.

Table 1--Estimates of Percent Net Undercount by Sex: 1940 to 2000
(a minus sign denotes a net overcount)

Category	Demographic Analysis							A.C.E.
	1940	1950	1960	1970	1980	1990	**2000**	**2000**
Total Population	6.5	4.1	3.1	2.7	1.2	1.8	-0.5	
Male	5.8	4.4	3.5	3.4	2.2	2.8	0.0	
Female	5.0	3.8	2.7	2.0	0.3	0.9	-1.1	
Male:Female Diff.	0.8	0.6	0.8	1.4	1.9	1.9	1.1	

Source: 1940-1990-- Robinson, J. Gregory, Bashir Ahmed, Prithwis Das Gupta, and Karen Woodrow, "Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis", Journal of the American Statistical Association, Vol. 88, No. 423, pp. 1061-1077. Estimates for 2000 are unpublished preliminary results.

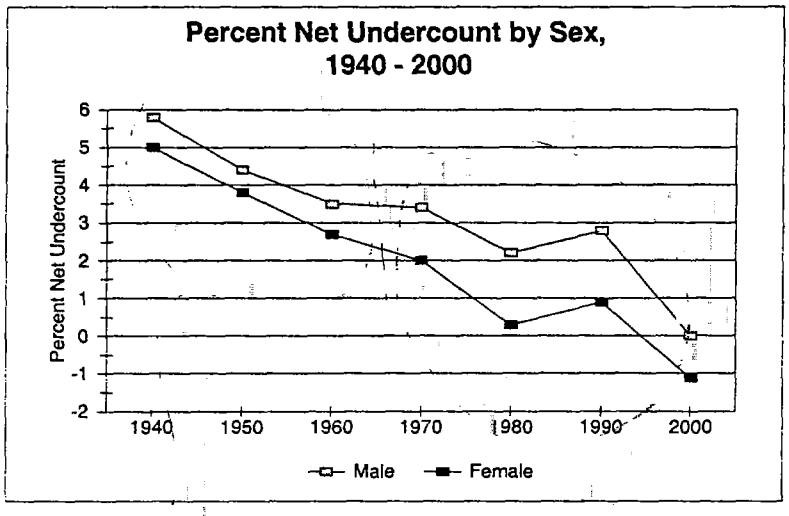


Table B--Comparison of Initial, Revised and 'Final' Estimates of Percent Net Undercount of the United States Resident Population Based on Demographic Analysis, by Sex and Race: 1980 and 1970

Sex and Race	1980				1970			
	1980 census cycle			1990 census cycle	1970 census cycle	1980 census cycle		1990 census cycle
	Initial estimate Feb. 1982 (b)	Revised estimate Aug. 1984 (c)	'Final' estimate Sept. 1985 (d)	Revised estimate April 1991 (e)	Published estimate Feb. 1974 (a)	Revised estimate Feb. 1982 (b)	'Final' estimate Sept. 1985 (d)	Revised estimate April 1991 (e)
Net undercount								
Total	-0.40	0.50	1.38	1.23	2.55	2.30	2.92	2.75
Male	0.50	1.50	2.37	2.19	3.30	3.10	3.70	3.51
Female	-1.20	-0.40	0.42	0.30	1.80	1.40	2.16	2.03
Black	4.80	5.30	5.89	4.53	7.70	7.60	8.02	6.60
Nonblack	-1.10	-0.20	0.75	0.77	1.80	1.50	2.24	2.25
Difference								
Male:female	1.70	1.90	1.95	1.89	1.50	1.70	1.54	1.48
Black:Nonblack	5.90	5.50	5.14	3.76	5.90	6.10	5.78	4.35

(a) U.S. Bureau of the Census, "Estimates of Coverage of Population by Sex, Race, and Age: Demographic Analysis," by Jacob S. Siegel. Evaluation and Research Program, PHC(E)-4, Feb. 1974. The estimates make no allowance for undocumented residents in 1970; the estimates represent coverage of the legally resident population.

(b) U.S. Bureau of the Census, "Coverage of the National Population in the 1980 Census, by Age, Sex, and Race: Preliminary Estimates by Demographic Analysis," by Jeffrey S. Passel, Jacob S. Siegel, and J. Gregory Robinson. Current Population Reports, Series P-23, No. 115, Feb. 1982. The demographic estimates (denominator of percent net undercount) make no allowance for undocumented residents; the 1980 census (numerator) is estimated to have included 2.06 million. This inconsistency is the major reason for the measured net overcounts shown in column 1.

(c) U.S. Bureau of the Census, "Revised Estimates of the Coverage of the Population in the 1980 Census Based on Demographic Analysis: A Report on Work in Progress," by Jeffrey S. Passel and J. Gregory Robinson. Paper Presented at the Annual Meeting of the American Statistical Association, Aug. 1984. The demographic estimates make no allowances for undocumented residents. The 1980 census is adjusted to exclude the 2.06 million "counted" undocumented aliens. Thus, the estimates represent coverage of the legally resident population.

(d) U.S. Bureau of the Census, "The Coverage of Population in the 1980 Census," by Robert Fay, Jeffrey S. Passel, and J. Gregory Robinson. Evaluation and Research Reports, PHC80-E4, Feb. 1988. The demographic estimates include allowance for 3 million undocumented residents; the estimates represent coverage of the total resident population.

(e) Revised estimates that are consistent with the preliminary 1990 Demographic Analysis estimates of net undercount for the total resident population.

Source: 1990 Decennial Census: Preliminary Research and Evaluation Memorandum No. 83, DA Evaluation Project D10, "Differences Between Preliminary and Final Estimates of Percent Net Undercount". by Bashir Ahmed and J. Gregory Robinson

Table 2--Estimates of Percent Net Undercount by Sex and Age: 1960 to 2000
(a minus sign denotes a net overcount)

Category	Demographic Analysis					A.C.E
	1960	1970	1980	1990	2000	2000
MALE						
Total	3.5	3.4	2.2	2.8	0.0	
0-17	2.8	2.7	0.9	2.2	-0.6	
18-29	5.9	3.9	3.3	2.2	-3.7	
30-49	4.2	5.1	3.6	3.8	1.4	
50+	2.2	2.5	1.2	2.7	1.3	
FEMALE						
Total	2.7	2.0	0.3	0.9	-1.1	
0-17	1.8	2.4	0.9	2.4	0.0	
18-29	2.8	1.3	0.4	0.6	-3.6	
30-49	1.9	1.3	-0.0	0.5	-0.8	
50+	4.6	2.6	-0.2	0.2	-0.8	

Note: DA estimates are consistent with estimates in Table 1.

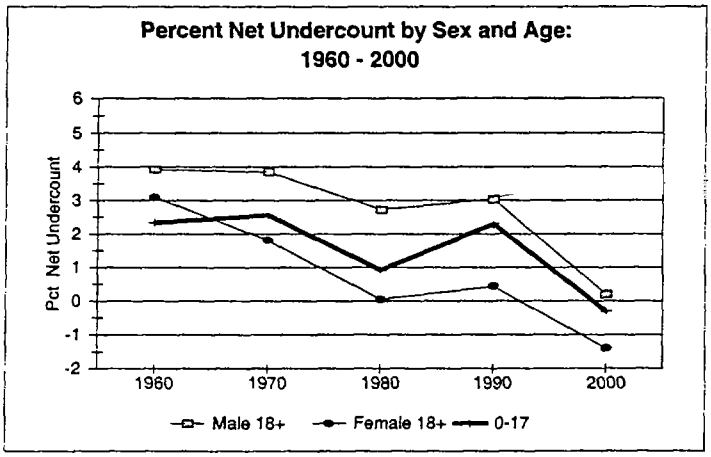
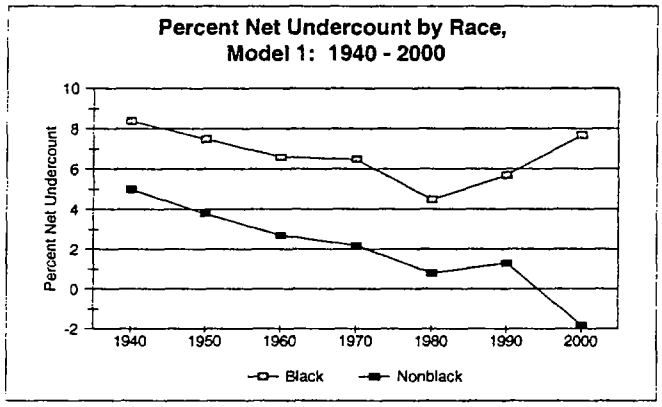


Table 3--Estimates of Percent Net Undercount by Race and Sex: 1940 to 2000
(a minus sign denotes a net overcount)

Category	Demographic Analysis								A.C.E.
							2000		
	1940	1950	1960	1970	1980	1990	Model 1	Model 2	2000
Total Population	6.5	4.1	3.1	2.7	1.2	1.8	-0.5	-0.5	
Black	8.4	7.5	6.6	6.5	4.5	5.7	7.7	3.0	
Male	10.9	9.7	8.8	9.1	7.5	8.5	10.1	5.5	
Female	6.0	5.4	4.4	4.0	1.7	3.0	5.4	0.7	
Nonblack	5.0	3.8	2.7	2.2	0.8	1.3	-1.8	-1.1	
Male	5.2	3.8	2.9	2.7	1.5	2.0	-1.5	-0.8	
Female	4.9	3.7	2.4	1.7	0.1	0.6	-2.1	-1.3	
Black:Nonblack Diff.	3.4	3.7	3.9	4.3	3.7	4.4	9.5	4.1	

Note: Model 1 census tabulations for Blacks include persons who marked the Black circle and no other race response
Model 2 census tabulations for Blacks include persons who marked the Black circle and other response circles.

Source: 1940-1990-- Robinson, J. Gregory, Bashir Ahmed, Prithwis Das Gupta, and Karen Woodrow, "Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis", Journal of the American Statistical Association, Vol. 88, No. 423, pp. 1061-1077. Estimates for 2000 are unpublished preliminary results.



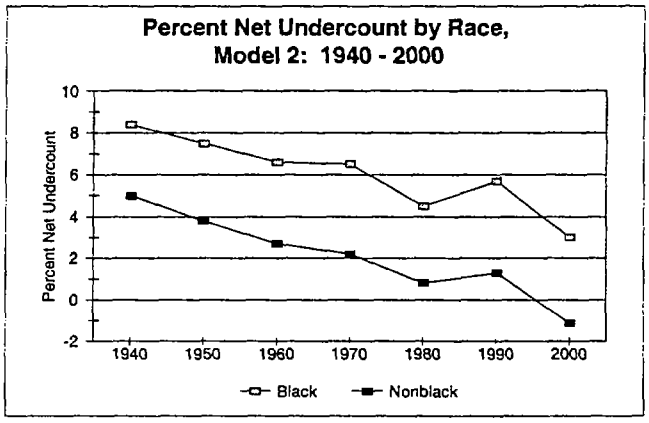
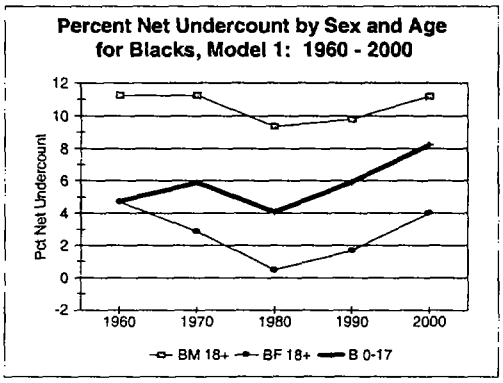


Table 4--Estimates of Percent Net Undercount by Race, Sex and Age: 1960 to 2000
(a minus sign denotes a net overcount)

	Demographic Analysis						A.C.E
					2000		
Category	1960	1970	1980	1990	Model 1	Model 2	2000
BLACK MALE							
Total	8.8	9.1	7.5	8.5	10.1	5.5	
0-17	5.4	6.2	4.2	5.9	7.9	-0.1	
18-29	15.1	12.1	9.2	7.7	10.9	6.8	
30-49	11.9	14.5	13.1	12.3	13.0	10.2	
50+	6.6	6.3	4.6	8.3	8.5	6.3	
BLACK FEMALE							
Total	4.4	4.0	1.7	3.0	5.4	0.7	
0-17	4.0	5.6	3.9	5.9	8.6	0.4	
18-29	5.4	4.5	2.4	2.9	5.3	0.6	
30-49	2.1	0.5	0.6	2.5	5.0	1.9	
50+	7.6	3.8	-1.9	-0.8	1.5	-0.8	
NONBLACK MALE							
Total	2.9	2.7	1.5	2.0	-1.5	-0.8	
0-17	2.4	2.1	0.3	1.5	-2.3	-0.7	
18-29	4.6	2.8	2.4	1.3	-6.3	-5.6	
30-49	3.4	4.0	2.5	2.7	-0.3	0.1	
50+	1.8	2.2	0.9	2.2	0.6	0.8	
NONBLACK FEMALE							
Total	2.4	1.7	0.1	0.6	-2.1	-1.3	
0-17	1.5	1.8	0.3	1.8	-1.8	-0.1	
18-29	2.4	0.9	0.1	0.3	-5.3	-4.4	
30-49	1.9	1.3	-0.1	0.2	-1.7	-1.2	
50+	4.3	2.5	-0.0	0.3	-1.1	-0.9	

Sources and notes: See Table 1 and 3



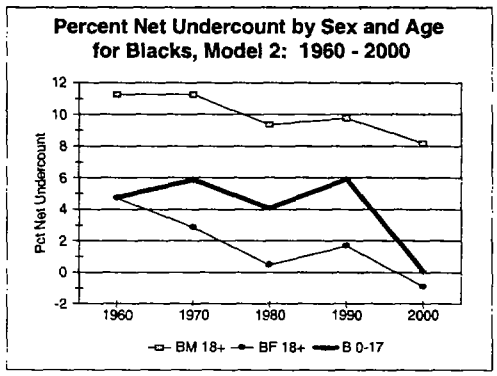
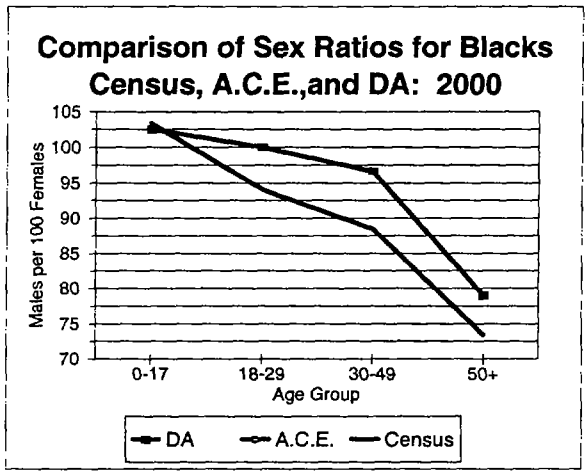


Table 5--Sex Ratios for the Census, Coverage Measurement Survey, and DA, by Race: 1990 and 2000

Category	1990			2000				
				DA	A.C.E.		Census	
	DA	PES	Census		Model 1	Model 2	Model 1	Model 2
BLACK								
Total	95.2	90.4	89.6	95.3			90.5	90.6
0-17	102.4	102.4	102.4	102.5			103.3	103.1
18-29	99.3	92.1	94.0	100.0			94.1	93.8
30-49	95.9	89.0	86.2	96.6			88.5	88.5
50+	78.3	72.1	71.5	79.0			73.4	73.4
NONBLACK								
Total	97.2	96.5	95.9	97.6			97.1	97.2
0-17	105.2	105.5	105.5	105.0			105.5	105.6
18-29	104.9	104.6	103.8	104.2			105.2	105.3
30-49	102.0	100.3	99.6	101.9			100.5	100.6
50+	80.8	79.9	79.4	84.5			83.1	83.1

Note: Model 1 census tabulations for Blacks includes persons who marked the Black circle and no other response circle. Model 2 census tabulations for Blacks includes persons who marked the Black circle and other response circles to the race question.

Source: DA, survey, and census data used to compute sex ratios are consistent with data used in Table 4.



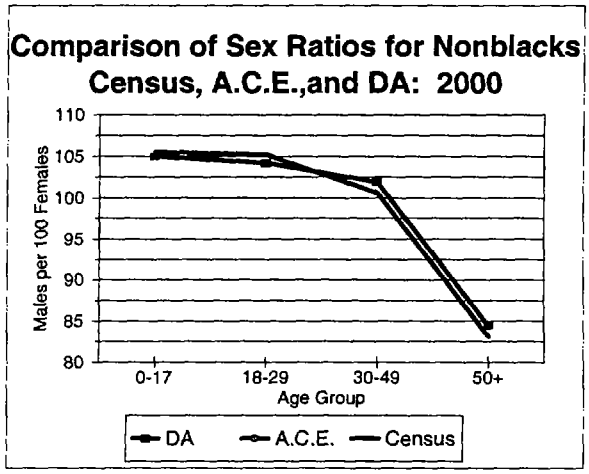
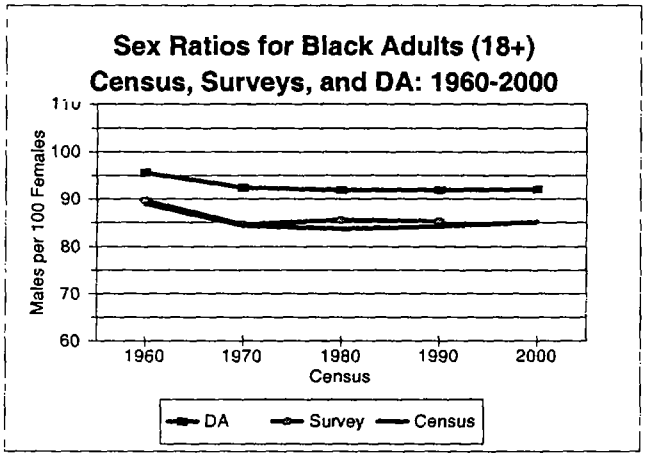


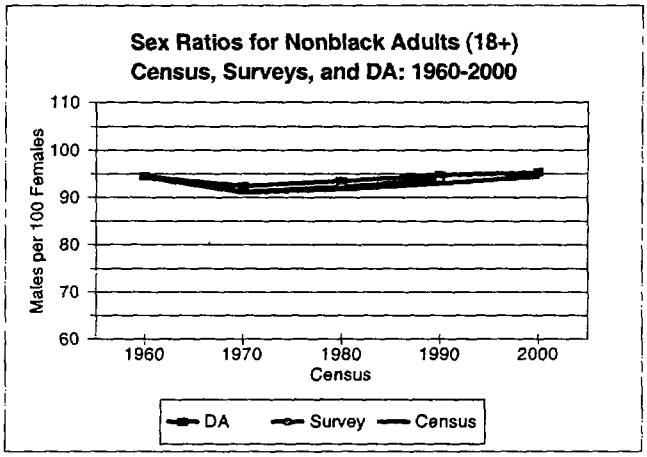
Table 6--Sex Ratios for the Census, Coverage Measurement Survey, and DA, for Adults by Race: 1960 - 2000

Category	1960	1970	1980	1990	2000 Model 1	2000 Model 2
BLACK						
DA	95.6	92.4	91.9	91.9	92.1	92.1
Survey	89.7	84.6	85.5	85.3		
Census	89.0	84.4	83.7	84.3	85.2	85.2
NONBLACK						
DA	94.5	92.4	93.5	94.7	95.3	95.3
Survey	94.4	91.2	92.3	93.7		
Census	94.3	91.1	91.7	92.9	94.5	94.6

Note: Model 1 census tabulations for Blacks includes persons who marked the Black circle and no other circle. Model 2 census tabulations for Blacks includes persons who marked the Black circle and other response circles to the race question.

DA and census sex ratios refer to the population 18+ in all years. Survey estimates are 18+ in 1990 (PES). For 1980 (PEP), coverage rates by sex for the population 20+ were assumed to represent coverage of the population 18+; for 1970 and 1960 the available survey undercount estimates for 15+ are used.





Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

12/13 ESCAP Presentation

- ★ **Universe=Hundred Percent Census Edited File (HCEF)**
- ★ **Nationwide, there are 115.9 million housing units**
- ★ **This is 0.4 percent more than expected ((HCEF - Hu. Est. April 2000)/Hu. Est. April 2000)**
- ★ **Counties that are at least 95% mailout/mailback have a shortage of 0.5 percent**
- ★ **Update/leave only counties have an overage of 5.1 percent**
- ★ **Counties with all other TEA combinations have an average overage of 1.2 percent**

Table 1. Difference Between the HCEF Housing Unit Count and the Housing Unit Estimate for the Nation and for Counties by Type of Enumeration Area

	National (1)	≥ 95% TEA=1 (2)	TEA=2 only (3)	TEA=1+2 only (4)	TEA=Mix (excl. 3+4) (5)	TEA=Mix (incl. 3+4) (6)
HCEF	115,915,874	63,465,541	5,443,378	28,888,644	12,090,894	6,027,417
Difference	506,843	-297,828	261,920	332,780	146,275	63,696
% Difference	0.4	-0.5	5.1	1.2	1.2	1.1
Number of Counties	3,142	391	818	1,301	415	217
Shortage						
HCEF	59,751,151	38,055,638	1,264,600	12,723,859	4,645,203	3,061,851
Difference	1,303,507	808,002	46,390	274,151	130,674	44,290
% Difference	2.1	2.1	3.5	2.1	2 7	2.0
Number of Counties	1,273	235	265	532	156	85
Overage						
HCEF	56,164,723	25,409,903	4,178,778	16,164,785	7,445,691	2,965,566
Difference	1,810,350	510,174	308,310	606,931	276,949	107,986
% Difference	3.3	2.1	8.0	3.9	3.9	4.0
Number of Counties	1,869	156	553	769	259	132

Types of Enumeration Area and Distribution of Counties by Type of Enumeration Area

In the Census 2000, there are nine types of enumeration areas:

TEA 1 - Block Canvassing and Mailout/Mailback
TEA 2 - Address Listing and Update/Leave
TEA 3 - List/Enumerate
TEA 4 - Remote Alaska
TEA 5 - "Rural" Update/Enumerate
TEA 6 - Military
TEA 7 - "Urban" Update/Leave
TEA 8 - "Urban" Update/Enumerate
TEA 9 - Additions to Address Listing Universe of Blocks

The distribution of the counties by type of enumeration areas:

Type of Enumeration Area	Number of Counties	Type of Enumeration Area	Number of Counties
1	**148**	5	13
2	**818**	5+8	1
1+2	**1,499**	5+8+9	1
1+9	11	5+6	1
1+8	1		**16**
1+7	7		
1+7+8	1	4	6
1+5+7	1	3	26
1+5+6+9	1	3+5	1
1+2+9	117	1+3	14
1+2+8	17	1+3+5+8	1
1+2+8+9	1	1+2+4	1
1+2+7	19	1+2+3	22
1+2+7+9	3	1+2+3+9	2
1+2+6	57	1+2+3+8	1
1+2+6+9	3	1+2+3+7+9	1
1+2+6+7	1	1+2+3+6	4
1+2+5	82	1+2+3+5	13
1+2+5+9	10	1+2+3+5+8	5
1+2+5+8	17	1+2+3+5+7	1
1+2+5+7	1	1+2+3+5+7+8	1
1+2+5+7+8	1	1+2+3+5+6	3
1+2+5+6	10	1+2+3+5+8	3
1+2+5+6+9	2	2+3	55
1+2+5+6+8	2	2+3+9	1
1+2+5+6+7+9	1	2+3+7+9	1
2+9	14	2+3+6	1
2+7+8+9	1	2+3+5	26
2+6	8	2+3+5+8	1
2+5	67	2+3+5+6	2
2+5+9	2	2+3+4	1
2+5+6	2	2+4	8
	460		**201**
TOTAL	**2,925**	TOTAL	**217**

Materials attached to these minutes were draft and preliminary material to inform the ESCAP Committee. The data and analysis contained in these documents are subject to revision and are not final. These materials report the results of research and analysis undertaken by Census Bureau staff. They have undergone a more limited review than official Census Bureau publications. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau.

Figure 1. 95%+ TEA1: Pct. Diff. HCEF-HU EST

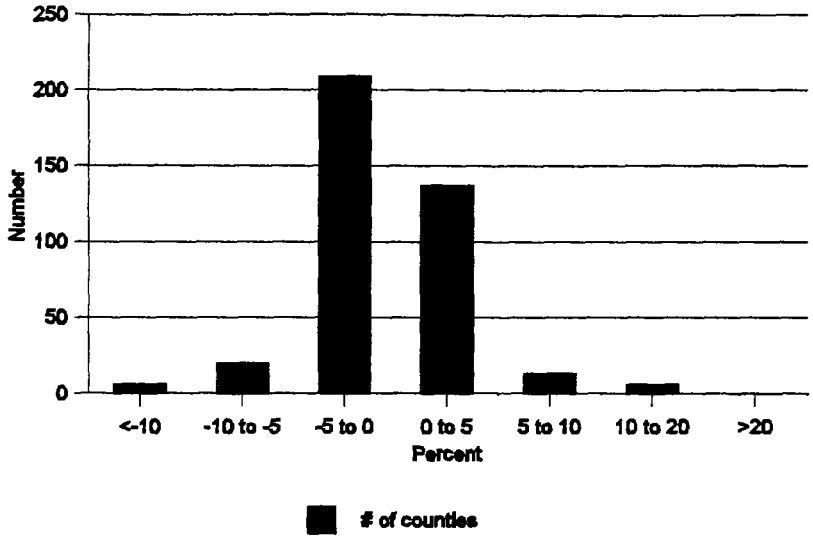


Figure 2. TEA2: Pct. Diff. HCEF-HU EST

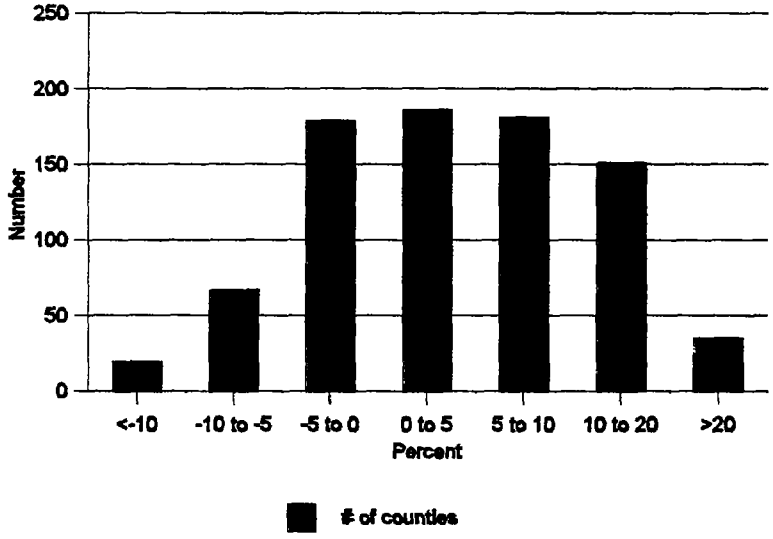


Figure 3. TEA1+TEA2: Pct. Diff. HCEF-HU EST

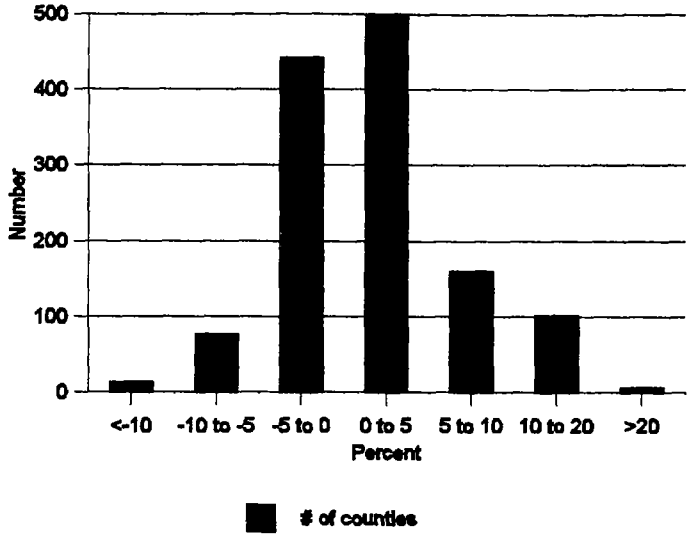


Figure 4. TEA-MIX (no 3 and 4): Pct. Diff. HCEF-HU EST

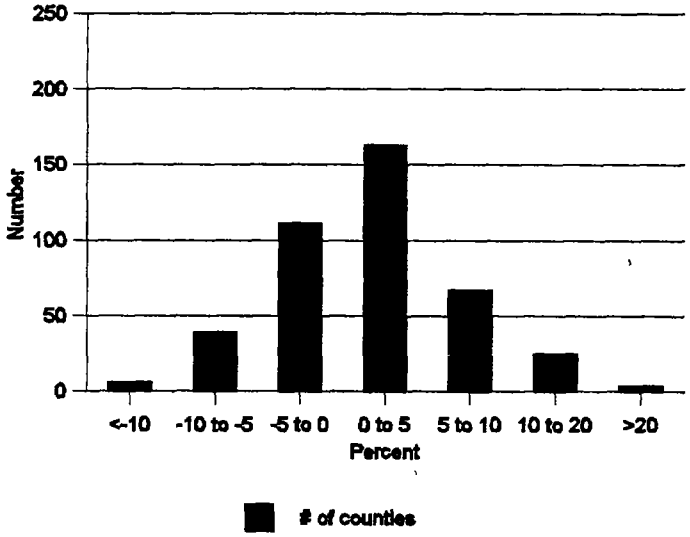
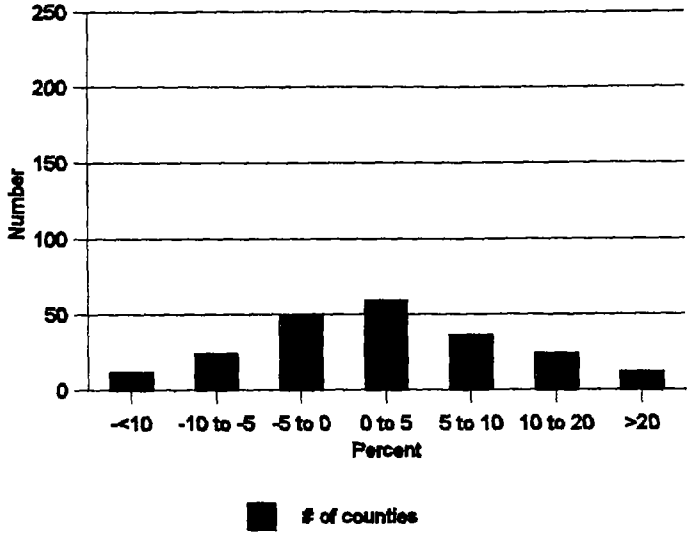


Figure 5. TEA-MIX (Incl. 3 & 4): Pct. Diff. HCEF-HU EST



ESCAP MEETING NO. 23 - 12/13/00

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 23
December 13, 2000

Prepared by: Nick Birnbaum

The twenty-third meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on December 13, 2000 at 10:30.

The agenda for the meeting was to briefly discuss the A.C.E. weight trimming procedure and to examine demographic analysis (DA) housing unit and population estimates compared to the final edited census file.

Committee Attendees:
William Barron
Nancy Potok
Paula Schneider
Cynthia Clark
Nancy Gordon
John Thompson
Jay Waite
Bob Fay
Howard Hogan
Ruth Ann Killion
John Long
Susan Miskura

Other Attendees:

Kenneth Prewitt	Roxie Jones
Tommy Wright	Kathleen Styles
Arjun Adlakha	Nick Birnbaum
Gregg Robinson	Carolee Bush
Kirsten West	Maria Urrutia
Signe Wetrogan	Annette Quinlan

I. Weight Trimming Procedure

Howard Hogan briefly discussed the weight trimming procedure for the A.C.E. In 1990, some sample block clusters had very high weights, thus having a disproportionate effect on the estimates. The A.C.E. was designed to limit how much the weights would vary. However, the Census Bureau specified a weight trimming procedure for handling any cluster whose weight exceeded a certain threshold as a further safeguard against disproportionate weights. One cluster in the United States and three in Puerto Rico did indeed marginally exceed the threshold, and consequently the weight trimming procedure was implemented on this cluster. It is important to note that cluster weighting does not appear to be a source of concern in the A.C.E. Additional information about the implementation of the weight trimming procedure will be presented at the next ESCAP meeting.

II. Demographic Analysis (DA) Housing Unit (HU) and Population Estimates

Gregg Robinson and Kirsten West presented DA population and housing unit estimates, respectively, as compared to the final edited census file.

Kirsten provided the same metrics as in her earlier ESCAP presentation (November 22, 2000) -- differences between the census file (in this case, the final edited file) housing unit counts and demographic benchmark analysis housing unit estimates for the nation and for counties by type of enumeration area (TEA). The results were discussed, and it was noted that there was general consistency between the two sets of numbers. The greatest percent difference was for counties that are solely update/leave. As mentioned at the November 22 meeting, this can be explained, in large part, by the greater uncertainty associated with the DA housing unit estimates for rural areas.

Gregg Robinson then presented various DA estimates including national population figures by sex, and DA estimates of percent net undercount in Census 2000 by race (Black and Nonblack), sex, and age group. Additionally, DA and Census 2000 sex ratios by age group were presented. These data were discussed, and it was noted that the use of multiple race reporting in Census 2000 must be taken into account when comparing demographic analysis estimates to Census 2000 tabulations of the Black and Nonblack populations. The results showed that differential coverage could be expected between the Black and Nonblack populations. However, the exact level of this differential was undetermined due to the effects of multiple race reporting.

III. Next Meeting

The agenda for the next meeting, to be held on December 20, is to examine the results of the weight trimming procedure and additional SBE data.